GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:54:52 ; Search time 159.389 Seconds
(without alignments)
268.450 Million cell updates/sec

Title: US-09-853-688-36
Perfect score: 19
Sequence: 1 tgtaggaagtctggggtgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	100.0	19	24	AAS18883	Growth hormone 1 g
c	2	19	100.0	3700	24	AAS18886	Growth hormone 1 g
	3	16.4	86.3	5333	22	AAS46366	Tumour suppressor
	4	16.4	86.3	5333	24	ABL32783	Human immune syste
c	5	15.8	83.2	65	24	ABN31023	Rat spliced transc
c	6	15.8	83.2	428	22	AAK54271	Murine transport a
	7	15.8	83.2	730	21	AAF08345	Fusarium venenatum
c	8	15.8	83.2	1612	22	AAD11871	Wild-type sheep BM
c	9	15.8	83.2	1612	22	AAD11872	Booroola sheep mut
c	10	15.8	83.2	2246	16	AAQ74064	Rat serotonin-1c c
c	11	15.8	83.2	2246	19	AAV48290	Rat 5-HT2C seroton
c	12	15.8	83.2	2246	19	AAV48291	Rat 5-HT2C S312F m
c	13	15.8	83.2	2246	19	AAV48288	Rat 5-HT2C seroton
c	14	15.8	83.2	2246	19	AAV48289	Rat 5-HT2C S312K m
c	15	15.8	83.2	2246	19	AAV48279	Rat 5-HT2C seroton
	16	15.8	83.2	6109	23	AAH78667	Murine Col5a3 cDNA
	17	15.8	83.2	10329	21	AAZ50264	Genomic DNA of pot
	18	15.8	83.2	10329	21	AAZ50653	Genomic DNA encodi
c	19	15.4	81.1	170	24	ABL77410	Human ovarian canc
c	20	15.4	81.1	456	22	ABA58575	Human foetal liver
c	21	15.4	81.1	456	22	AAK06696	Human brain expres
c	22	15.4	81.1	456	22	AAK32395	Human bone marrow
c	23	15.4	81.1	456	22	AAI38237	Probe #6923 used t
c	24	15.4	81.1	456	24	ABS07178	Human genome-deriv
	25	15.4	81.1	700	22	AAH92715	Human inflammatory
	26	15.4	81.1	881	24	ABQ46242	Oligonucleotide fo
c	27	15.4	81.1	881	24	ABQ46243	Oligonucleotide fo
	28	15.4	81.1	901	24	ABQ33436	Oligonucleotide fo
c	29	15.4	81.1	901	24	ABQ33437	Oligonucleotide fo
	30	15.4	81.1	1343	21	AAC59837	Human secreted pro
c	31	15.4	81.1	4038	10	AAN90382	Genes encoding hum
c	32	15.4	81.1	34917	22	AAK70686	Human immune/haema
	33	15	78.9	392	24	ABK62503	Rat sequence diffe
c	34	15	78.9	2207	24	ABK63542	Rat sequence diffe
c	35	15	78.9	24757	22	AAS27687	DNA encoding novel
	36	15	78.9	24757	22	AAS33481	DNA encoding human
c	37	14.8	77.9	393	21	AAF15743	Human prostate can
c	38	14.8	77.9	393	22	AAS00856	Human cDNA clone H
c	39	14.8	77.9	558	22	AAK36046	Human bone marrow
c	40	14.8	77.9	558	22	AAI41762	Probe #10448 used
	41	14.8	77.9	640	22	AAH71209	Human cervical can
c	42	14.8	77.9	653	21	AAC75146	Human ORFX ORF701
c	43	14.8	77.9	653	24	ABN25315	Human ORFX polynuc
c	44	14.8	77.9	698	22	AAF25821	C. glutamicum guan
c	45	14.8	77.9	787	23	AAS71701	DNA encoding novel

ALIGNMENTS

RESULT 1
AAS18883
ID AAS18883 standard; DNA; 19 BP.
XX
AC AAS18883;
XX
DT 12-MAR-2002 (first entry)
XX
DE Growth hormone 1 gene (GH1) specific fragment, PCR primer GH1R.
XX
KW Growth hormone 1; GH1; osteopathic; gene therapy; protein therapy;
KW diabetes; obesity; infection; acromegaly; gigantism; sodium retention;
KW water retention; metabolic syndrome; mood disorder; sleep disorder;
KW Growth hormone dysfunction; familial growth hormone deficiency;
KW short stature; pituitary storage defect; human; PCR primer; GH1R; ss.
XX
OS Homo sapiens.
XX
PN WO200185993-A2.
XX
PD 15-NOV-2001.
XX
PF 14-MAY-2001; 2001WO-GB02126.
XX
PR 12-MAY-2000; 2000GB-0011459.
PR 14-JUL-2000; 2000EP-0306004.
XX
PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
XX
PI Cooper DN, Procter AM, Gregory J, Millar DS;
XX
DR WPI; 2002-089798/12.

```
XX
PT   Detecting growth hormone variants (GH1), useful in screening patients
PT   for growth hormone irregularities, comprises comparing the nucleotide
PT   sequence of a GH1 gene from a test sample with that of a standard
PT   sequence of the human GH1 -
XX
PS   Example 2; Page 39; 95pp; English.
XX
CC   The invention described a method of detecting variation in growth hormone
CC   1 (GH1), and therefore GH dysfunction in an individual. The method
CC   comprises comparing the nucleotide sequence of GH1 gene obtained from the
CC   test sample with a standard human GH1 gene sequence, in order to identify
CC   variation (GH1 variant). The method is useful in screening patients for
CC   growth hormone irregularities or producing variant proteins for treating
CC   irregularities, and for the early detection and appropriate clinical
CC   management of familial GH deficiency. The GH1 variants are useful in
CC   therapeutic, diagnostic or detection method, particularly for determining
CC   binding defects and susceptibility to a disease such as diabetes, obesity
CC   or infection; for treating acromegaly or gigantism conditions associated
CC   with lactogenic, diabetogenic, lipolytic and protein anabolic effects,
CC   conditions associated with sodium and water retention, metabolic
CC   syndromes, mood and sleep disorders; diagnosing GH dysfunction and
CC   determining pituitary storage defects. The GH1 variants are especially
CC   useful in gene therapy or protein therapy. The GH1 or GH variant may also
CC   be used in the preparation of a medicament, diagnostics composition or
CC   kit, or detection kit. The method has the advantage of: expanding the
CC   know spectrum of GH1 gene mutations; evaluating the role of GH1 gene
CC   mutations in the etiology of short stature; identifying of the mode of
CC   inheritance of novel lesions; evaluation the effects of GH1 mutations on
CC   the structure and function of the GH molecule and development of rapid
CC   diagnostic tests for inherited GH deficiency. This sequence is the GH1
CC   PCR primer, GH1R, used with GH1F (AAS18882) to amplify a GH1-specific
CC   fragment, described in the method of the invention.
XX
SQ   Sequence 19 BP; 3 A; 2 C; 9 G; 5 T; 0 other;

  Query Match             100.0%;  Score 19;  DB 24;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||||||||
Db       1 TGTAGGAAGTCTGGGGTGC 19


RESULT 2
AAS18886/c
ID   AAS18886 standard; DNA; 3700 BP.
XX
AC   AAS18886;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Growth hormone 1 gene (GH1), reference sequence.
XX
KW   Growth hormone 1; GH1; osteopathic; gene therapy; protein therapy;
KW   diabetes; obesity; infection; acromegaly; gigantism; sodium retention;
KW   water retention; metabolic syndrome; mood disorder; sleep disorder;
KW   Growth hormone dysfunction; familial growth hormone deficiency;
KW   short stature; pituitary storage defect; human; chromosome 17q23; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     763..1100
FT                   /*tag= a
FT   CDS             763..2230
FT                   /*tag= b
FT                   /product= "GH1"
FT                   /note= "Growth hormone 1"
FT   variation       replace(124,G)
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(128,T)
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(134,G)
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(135,T)
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(135,C)
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(136,G)
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(141,G)
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(179,C)
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(188,T)
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(218,A)
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(226,G)
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(230,C)
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(234,C)
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(236,C)
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(249,G)
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(281,C)
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(284,A)
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(284,C)
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(286,C)
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(303,C)
FT                   /*tag= v
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(313,A)
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(508)
FT                   /*tag= x
FT                   /note= "Deletion of base A"
FT   variation       replace(519,T)
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(524,A)
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(558,G)
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(565,G)
```

```
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(573,G)
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(580,G)
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(585,G)
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(586,T)
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(620,A)
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(622,G)
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(649,C)
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(665,C)
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(670,G)
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(676,A)
FT                   /*tag= al
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(685,A)
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            701..772
FT                   /*tag= an
FT                   /number= 1
FT   intron          773..1032
FT                   /*tag= ao
FT                   /number= 1
FT   variation       replace(836,C)
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(839,C)
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(879,G)
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(883,A)
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(901,C)
FT                   /*tag= at
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1010,T)
FT                   /*tag= au
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            1033..1193
FT                   /*tag= av
FT                   /number= 2
FT   variation       replace(1097,A)
FT                   /*tag= aw
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1101,T)
FT                   /*tag= ax
FT                   /standard_name= "Single nucleotide polymorphism"
FT   mat_peptide     1101..2227
FT                   /*tag= ay
FT                   /label= "mature_GH1"
FT   variation       replace(1114,T)
FT                   /*tag= az
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1169,A)
FT                   /*tag= ba
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1182,T)
FT                   /*tag= bb
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1189,G)
FT                   /*tag= bc
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1193,G)
FT                   /*tag= bd
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          1194..1402
FT                   /*tag= be
FT                   /number= 2
FT   variation       replace(1196,G)
FT                   /*tag= bf
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1196,C)
FT                   /*tag= bg
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1208,C)
FT                   /*tag= bh
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1212,T)
FT                   /*tag= bi
FT                   /standard_name= "Single nucleotide polymorphism"
```

```
  Query Match             100.0%;  Score 19;  DB 24;  Length 3700;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||||||||
Db    3316 TGTAGGAAGTCTGGGGTGC 3298


RESULT 3
AAS46366
ID   AAS46366 standard; DNA; 5333 BP.
XX
AC   AAS46366;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Tumour suppressor gene derived chemically modified sequence #88.
XX
KW   Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW   cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW   cytosine methylation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200168912-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02955.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-602752/68.
XX
PT   Fragments of chemically modified genes associated with tumour suppressor
```

```
PT   genes and oncogenes, useful in designing primers and probes for
PT   analysing diseases associated with cytosine methylation state e.g.
PT   cancer -
XX
PS   Claim 1; SEQ ID No 88; 27pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of 18
CC   bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC   bisulphite, of genes associated with tumour suppression and
CC   oncogenes having a sequence taken from 536 (actually 533 since
CC   numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC   (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC   peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC   form part of a set of probes for detecting the cytosine methylation state
CC   and/or single nucleotide polymorphisms and also to be used in an
CC   array for analysing diseases associated with CpG dinucleotides e.g.
CC   cancers and tumours. The probes can also be used in a method for
CC   ascertaining genetic and/or epigenetic parameters for the diagnosis
CC   and/or therapy of existing diseases or the predisposition to specific
CC   diseases, by analysing cytosine methylations. The parameters may be
CC   compared to another set of genetic and/or epigenetic parameters, the
CC   differences serving as basis for diagnosis and/or prognosis events which
CC   are disadvantageous to patients. The present sequence is one of the
CC   533 genomic sequences derived from tumour suppressor genes and
CC   oncogenes. Sequences with even numbered Seq ID numbers are the
CC   complementary sequence of the corresponding odd numbered sequence (e.g.
CC   ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC   is missing).
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5333 BP; 1275 A; 101 C; 1345 G; 2612 T; 0 other;

  Query Match             86.3%;  Score 16.4;  DB 22;  Length 5333;
  Best Local Similarity   94.4%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           |||||||||| |||||||
Db    4336 TGTAGGAAGTTTGGGGTG 4353


RESULT 4
ABL32783
ID   ABL32783 standard; DNA; 5333 BP.
XX
AC   ABL32783;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 756.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07537.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation -
XX
PS   Claim 1; SEQ ID NO 756; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 5333 BP; 1275 A; 101 C; 1345 G; 2612 T; 0 other;

  Query Match             86.3%;  Score 16.4;  DB 24;  Length 5333;
  Best Local Similarity   94.4%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           |||||||||| |||||||
Db    4336 TGTAGGAAGTTTGGGGTG 4353


RESULT 5
ABN31023/c
ID   ABN31023 standard; DNA; 65 BP.
XX
AC   ABN31023;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:3771.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 3771; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
```

```
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 12 C; 21 G; 13 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 24;  Length 65;
  Best Local Similarity   89.5%;  Pred. No. 2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||| | |||
Db      53 TGTAGGAAGTCTGCGCTGC 35


RESULT 6
AAK54271/c
ID   AAK54271 standard; cDNA; 428 BP.
XX
AC   AAK54271;
XX
DT   16-NOV-2001  (first entry)
XX
DE   Murine transport and binding associated protein encoding cDNA SEQ ID 836.
XX
KW   Murine; liver; gene library; amino acid synthesis; binding protein;
KW   cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
KW   phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
KW   replication; transcription; translation; transport protein; ss.
XX
OS   Mus musculus.
XX
PN   DE20103510-U1.
XX
PD   07-JUN-2001.
XX
PF   28-FEB-2001; 2001DE-2003510.
XX
PR   02-DEC-1999;   99DE-1058160.
XX
PA   (LION-) LION BIOSCIENCE AG.
XX
DR   WPI; 2001-368570/39.
XX
PT   Gene library containing sequences with specific 3'-ends and no polyA
PT   tail, encoding proteins involved in a wide range of cellular processes
PT   -
XX
PS   Claim 15; Page 250; 251pp; German.
XX
CC   This invention describes a novel gene library (A) comprises a gene
CC   sequence (or its part) encoding a protein involved in amino acid
CC   synthesis, cellular/energy metabolism, metabolism of
CC   fatty acids/phospholipids, synthesis or breakdown of
CC   purines/pyrimidines/nucleosides/nucleotides, DNA
CC   replication/transcription/translation, or is a transport/binding protein.
CC   (A) are produced that correspond to the 3'-end of mRNA but without the
CC   polyA tail. They can be prepared more efficiently and with less effort
CC   than conventional libraries. AAK53436-AAK54275 represent fragments of the
CC   gene library described in the method of the invention.
XX
SQ   Sequence 428 BP; 108 A; 99 C; 122 G; 99 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 22;  Length 428;
  Best Local Similarity   89.5%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           || |||||||||||||| |
Db     126 TGAAGGAAGTCTGGGGTAC 108


RESULT 7
AAF08345
ID   AAF08345 standard; cDNA; 730 BP.
XX
AC   AAF08345;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Fusarium venenatum EST SEQ ID NO:868.
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Fusarium venenatum.
XX
PN   WO200056762-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US07781.
XX
PR   22-MAR-1999;   99US-0273623.
XX
PA   (NOVO ) NOVO NORDISK BIOTECH INC.
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Berka RM,  Key MW,  Shuster JR,  Kauppinen S,  Clausen IG,  Olsen PB;
XX
DR   WPI; 2000-594572/56.
XX
PT   Monitoring differential expression of genes in filamentous fungal cells
PT   uses fluorescence-labeled nucleic acids isolated from the cells and a
PT   substrate of expressed sequence tags -
XX
PS   Claim 86; Page 715; 3161pp; English.
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labeled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring
CC   the global expression of genes from FF cells allows the production
CC   potential of the microorganisms to be improved. New genes may be
CC   discovered, possible functions of unknown open reading frames can be
CC   identified and gene copy number variation and stability can be
CC   monitored. The expression of genes can be used to study how FF cells
CC   adapt to changes in culture conditions, environmental stress, spore
CC   morphogenesis, recombination, metabolic or catabolic pathway
CC   engineering. Using ESTs provides several advantages over genomic or
CC   random cDNA clones including elimination of redundancy as one spot on an
```

```
CC   array equals one gene or open reading frame, and organisation of the
CC   microarrays based on function of the gene products to facilitate
CC   analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC   Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC   niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC   AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC   all specifically claimed in the present invention.
XX
SQ   Sequence 730 BP; 205 A; 166 C; 192 G; 156 T; 11 other;

  Query Match             83.2%;  Score 15.8;  DB 21;  Length 730;
  Best Local Similarity   89.5%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           || |||||||||||||| |
Db     280 TGAAGGAAGTCTGGGGTTC 298


RESULT 8
AAD11871/c
ID   AAD11871 standard; cDNA; 1612 BP.
XX
AC   AAD11871;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Wild-type sheep BMP1B receptor cDNA.
XX
KW   Bone morphogenetic protein 1B receptor; BMP1B receptor; ovulation;
KW   FecB; fecundity; fertility; Booroola; contraceptive;
KW   transforming growth factor-beta family; chromosome 6; sheep; ss.
XX
OS   Ovis aries.
XX
FH   Key             Location/Qualifiers
FT   promoter        1..84
FT                   /*tag= a
FT   CDS             85..1593
FT                   /*tag= b
FT                   /product= "Wild-type BMP1B receptor"
XX
PN   WO200148204-A1.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-NZ00259.
XX
PR   23-DEC-1999;   99NZ-0502058.
XX
PA   (AGRE-) AGRESEARCH LTD.
XX
PI   Wilson TM,  Wu X;
XX
DR   WPI; 2001-441717/47.
DR   P-PSDB; AAE06221.
XX
PT   Novel isolated bone morphogenetic protein 1B receptor polypeptide
PT   useful for modulating the ovulation rate of a female vertebrate -
XX
PS   Example; Fig 2; 53pp; English.
XX
CC   The invention relates to a bone morphogenetic protein 1B (BMP1B)
CC   receptor polypeptide comprising a sequence which differs from that
CC   of the wild type in that residue 249 is arginine and not glutamine,
CC   or a sequence in which residue 249 is glutamine, but which is otherwise
CC   different from the wild type BMP1B polypeptide sequence and which has
CC   the ability to modulate ovulation in a female mammal. Mutation in
CC   the BMP1B receptor gene is responsible for increased ovulation rate
CC   in sheeps derived from Booroola Merino strain that carry an autosomal
CC   mutation in FecB/Booroola gene. The FecB gene is mapped to chromosome 6.
CC   The BMP1B receptor of the invention and the polynucleotide encoding
CC   it are useful for modulating the ovulation rate of a female vertebrate.
CC   Identification of mutated BMP1B receptor nucleic acid molecule in a
CC   vertebrate, is useful for assessing fecundity in vertebrate such as
CC   humans and other commercially important mammals and birds including
CC   sheep, cattle, horses, goats, deer, pigs, cats, dogs, possums, and
CC   poultry. The polypeptide is useful to raise antibodies and for reducing
CC   unwanted populations of feral vertebrates. The polynucleotide is useful
CC   for identifying sequence variants in individual animals that are
CC   associated with increased ovulation. The present sequence is
CC   BMP1B receptor cDNA from wild-type sheep. The BMP1B receptor is a
CC   member of transforming growth factor-beta family.
XX
SQ   Sequence 1612 BP; 473 A; 359 C; 394 G; 386 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 22;  Length 1612;
  Best Local Similarity   89.5%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           || ||||||| ||||||||
Db     420 TGGAGGAAGTGTGGGGTGC 402


RESULT 9
AAD11872/c
ID   AAD11872 standard; cDNA; 1612 BP.
XX
AC   AAD11872;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Booroola sheep mutated BMP1B receptor cDNA.
XX
KW   Bone morphogenetic protein 1B receptor; BMP1B receptor; ovulation;
KW   FecB; fecundity; fertility; Booroola; contraceptive; mutant; mutein;
KW   transforming growth factor-beta family; chromosome 6; sheep; ss.
XX
OS   Ovis aries.
XX
FH   Key             Location/Qualifiers
FT   promoter        1..84
FT                   /*tag= a
FT   CDS             85..1572
FT                   /*tag= b
FT                   /product= "Mutated BMP1B receptor"
FT                   /note= "Does not include stop codon"
FT                   /partial
FT   mutation        replace (830, A)
FT                   /*tag= c
XX
PN   WO200148204-A1.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-NZ00259.
XX
PR   23-DEC-1999;   99NZ-0502058.
XX
PA   (AGRE-) AGRESEARCH LTD.
XX
PI   Wilson TM,  Wu X;
XX
DR   WPI; 2001-441717/47.
DR   P-PSDB; AAE06222.
XX
PT   Novel isolated bone morphogenetic protein 1B receptor polypeptide
PT   useful for modulating the ovulation rate of a female vertebrate -
XX
PS   Claim 3; Page 48-50; 53pp; English.
XX
CC   The invention relates to a bone morphogenetic protein 1B (BMP1B)
CC   receptor polypeptide comprising a sequence which differs from that
CC   of the wild type in that residue 249 is arginine and not glutamine,
CC   or a sequence in which residue 249 is glutamine, but which is otherwise
```

```
CC   different from the wild type BMP1B polypeptide sequence and which has
CC   the ability to modulate ovulation in a female mammal. Mutation in
CC   the BMP1B receptor gene is responsible for increased ovulation rate
CC   in sheeps derived from Booroola Merino strain that carry an autosomal
CC   mutation in FecB/Booroola gene. The FecB gene is mapped to chromosome 6.
CC   The BMP1B receptor of the invention and the polynucleotide encoding
CC   it are useful for modulating the ovulation rate of a female vertebrate.
CC   Identification of mutated BMP1B receptor nucleic acid molecule in a
CC   vertebrate, is useful for assessing fecundity in vertebrate such as
CC   humans and other commercially important mammals and birds including
CC   sheep, cattle, horses, goats, deer, pigs, cats, dogs, possums, and
CC   poultry. The polypeptide is useful to raise antibodies and for reducing
CC   unwanted populations of feral vertebrates. The polynucleotide is useful
CC   for identifying sequence variants in individual animals that are
CC   associated with increased ovulation. The present sequence is mutated
CC   BMP1B receptor cDNA from Booroola sheep. The BMP1B receptor is a
CC   member of transforming growth factor-beta family.
XX
SQ   Sequence 1612 BP; 472 A; 359 C; 395 G; 386 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 22;  Length 1612;
  Best Local Similarity   89.5%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           || ||||||| ||||||||
Db     420 TGGAGGAAGTGTGGGGTGC 402


RESULT 10
AAQ74064/c
ID   AAQ74064 standard; cDNA to mRNA; 2246 BP.
XX
AC   AAQ74064;
XX
DT   29-JAN-1996  (first entry)
XX
DE   Rat serotonin-1c cDNA.
XX
KW   Serotonin; primer; mRNA; specificity; pharmaceutical; ss.
XX
OS   Rattus rattus.
XX
PN   JP07123984-A.
XX
PD   16-MAY-1995.
XX
PF   05-NOV-1993;   93JP-0275852.
XX
PR   05-NOV-1993;   93JP-0275852.
XX
PA   (HITB ) HITACHI CHEM CO LTD.
XX
DR   WPI; 1995-211627/28.
XX
PT   A primer for the detection and the determn. of a specific messenger
PT   RNA - can detect and determine specific mRNA(s) with high
PT   reliability
XX
PS   Example 28; Page 30-31; 35pp; Japanese.
XX
CC   AAQ74064 is rat serotonin 1c cDNA. This cDNA is amplified by the
CC   primers AAQ74047 and AAQ74048. The primers are used specifically for the
CC   detection and isolation of this sequence. They have the advantage of
CC   high sensitivity and reliability and are useful in the pharmaceutical
CC   industry.
XX
SQ   Sequence 2246 BP; 532 A; 542 C; 544 G; 628 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 16;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 11
AAV48290/c
ID   AAV48290 standard; DNA; 2246 BP.
XX
AC   AAV48290;
XX
DT   21-JAN-1999  (first entry)
XX
DE   Rat 5-HT2C serotonin receptor S312F mutant DNA.
XX
KW   Rat 5-HT2C serotonin receptor; G-protein-coupled receptor;
KW   constitutively activated monoamine G protein-coupled receptor;
KW   screen; agonist; inverse agonist; antagonist; mutant; ss.
XX
OS   Synthetic.
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             688..2070
FT                   /*tag= a
FT                   /product= Mutant 5-HT2C serotonin receptor
FT   mutation        1621..1623
FT                   /*tag= b
FT                   /note= "codon has been changed from TCC to TTC"
XX
PN   WO9838217-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-US03991.
XX
PR   07-OCT-1997;   97US-0061268.
PR   27-FEB-1997;   97US-0039465.
XX
PA   (EGAN/) EGAN C C.
PA   (HERR/) HERRICK-DAVIS K.
PA   (TEIT/) TEITLER M.
XX
PI   Egan CC,  Herrick-Davis K,  Teitler M;
XX
DR   WPI; 1998-495389/42.
DR   P-PSDB; AAW77113.
XX
PT   Method of constitutively activating targeted G-protein coupled
PT   monoamine receptor - comprises use of site directed mutagenesis,
PT   useful for, e.g. screening for agonists and antagonists of native
PT   receptor
XX
PS   Example 2; Fig 35; 97pp; English.
XX
CC   The present sequence encodes a rat 5-HT2C serotonin receptor S312F
CC   mutant. 5-HT2C is a G-protein-coupled receptor. The sequence of the
CC   wild type receptor (AAV48279) is modified using the method of the
CC   invention to produce the present sequence. The mutant receptor exists
CC   in a constitutively activated state exhibiting both a greater response
CC   to agonists and a coupling to the G protein second messenger system
CC   even in the absence of agonists. The constitutively activated monoamine
CC   G protein-coupled receptor can be used to screen for agonists, inverse
CC   agonists, and antagonists of the native receptor.
XX
SQ   Sequence 2246 BP; 537 A; 542 C; 538 G; 629 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 19;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||| | |||
```

```
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 12
AAV48291/c
ID   AAV48291 standard; DNA; 2246 BP.
XX
AC   AAV48291;
XX
DT   21-JAN-1999  (first entry)
XX
DE   Rat 5-HT2C S312F mutant receptor DNA with a unique restriction site.
XX
KW   Rat 5-HT2C serotonin receptor; G-protein-coupled receptor;
KW   constitutively activated monoamine G protein-coupled receptor;
KW   screen; agonist; inverse agonist; antagonist; mutant; ss.
XX
OS   Synthetic.
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             688..2070
FT                   /*tag= a
FT                   /product= Mutant 5-HT2C serotonin receptor
FT   mutation        1621..1623
FT                   /*tag= b
FT                   /note= "codon has been changed from TCC to TTC"
FT   mutation        1629
FT                   /*tag= c
FT                   /note= "C changed to A to create a ScaI site"
XX
PN   WO9838217-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-US03991.
XX
PR   07-OCT-1997;   97US-0061268.
PR   27-FEB-1997;   97US-0039465.
XX
PA   (EGAN/) EGAN C C.
PA   (HERR/) HERRICK-DAVIS K.
PA   (TEIT/) TEITLER M.
XX
PI   Egan CC,  Herrick-Davis K,  Teitler M;
XX
DR   WPI; 1998-495389/42.
XX
PT   Method of constitutively activating targeted G-protein coupled
PT   mono:amine receptor - comprises use of site directed mutagenesis,
PT   useful for, e.g. screening for agonists and antagonists of native
PT   receptor
XX
PS   Claim 20; Fig 36; 97pp; English.
XX
CC   The present sequence encodes a rat 5-HT2C serotonin receptor S312F
CC   mutant. The present sequence contains a ScaI restriction site. 5-HT2C
CC   is a G-protein-coupled receptor. The sequence of the wild type receptor
CC   (AAV48279) is modified using the method of the invention to produce the
CC   present sequence. The mutant receptor exists in a constitutively
CC   activated state exhibiting both a greater response to agonists and a
CC   coupling to the G protein second messenger system even in the absence
CC   of agonists. The constitutively activated monoamine G protein-coupled
CC   receptor can be used to screen for agonists, inverse agonists, and
CC   antagonists of the native receptor.
XX
SQ   Sequence 2246 BP; 538 A; 541 C; 538 G; 629 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 19;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGTAGGAAGTCTGGGGTGC 19
          ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 13
AAV48288/c
ID   AAV48288 standard; DNA; 2246 BP.
XX
AC   AAV48288;
XX
DT   21-JAN-1999  (first entry)
XX
DE   Rat 5-HT2C serotonin receptor S312K mutant DNA.
XX
KW   Rat 5-HT2C serotonin receptor; G-protein-coupled receptor;
KW   constitutively activated monoamine G protein-coupled receptor;
KW   screen; agonist; inverse agonist; antagonist; mutant; ss.
XX
OS   Synthetic.
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             688..2070
FT                   /*tag= a
FT                   /product= Mutant 5-HT2C serotonin receptor
FT   mutation        1621..1623
FT                   /*tag= b
FT                   /note= "codon has been changed from TCC to AAG"
XX
PN   WO9838217-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-US03991.
XX
PR   07-OCT-1997;   97US-0061268.
PR   27-FEB-1997;   97US-0039465.
XX
PA   (EGAN/) EGAN C C.
PA   (HERR/) HERRICK-DAVIS K.
PA   (TEIT/) TEITLER M.
XX
PI   Egan CC,  Herrick-Davis K,  Teitler M;
XX
DR   WPI; 1998-495389/42.
DR   P-PSDB; AAW77112.
XX
PT   Method of constitutively activating targeted G-protein coupled
PT   mono:amine receptor - comprises use of site directed mutagenesis,
PT   useful for, e.g. screening for agonists and antagonists of native
PT   receptor
XX
PS   Example 2; Fig 32; 97pp; English.
XX
CC   The present sequence encodes a rat 5-HT2C serotonin receptor S312K
CC   mutant. 5-HT2C is a G-protein-coupled receptor. The sequence of the
CC   wild type receptor (AAV48279) is modified using the method of the
CC   invention to produce the present sequence. The mutant receptor exists
CC   in a constitutively activated state exhibiting both a greater response
CC   to agonists and a coupling to the G protein second messenger system
CC   even in the absence of agonists. The constitutively activated monoamine
CC   G protein-coupled receptor can be used to screen for agonists, inverse
CC   agonists, and antagonists of the native receptor.
XX
SQ   Sequence 2246 BP; 539 A; 541 C; 539 G; 627 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 19;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGTAGGAAGTCTGGGGTGC 19
          ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078
```

```
RESULT 14
AAV48289/c
ID   AAV48289 standard; DNA; 2246 BP.
XX
AC   AAV48289;
XX
DT   21-JAN-1999  (first entry)
XX
DE   Rat 5-HT2C S312K mutant receptor DNA with a unique restriction site.
XX
KW   Rat 5-HT2C serotonin receptor; G-protein-coupled receptor;
KW   constitutively activated monoamine G protein-coupled receptor;
KW   screen; agonist; inverse agonist; antagonist; mutant; ss.
XX
OS   Synthetic.
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             688..2070
FT                   /*tag= a
FT                   /product= Mutant 5-HT2C serotonin receptor
FT   mutation        1621..1623
FT                   /*tag= b
FT                   /note= "codon has been changed from TCC to AAG"
FT   mutation        1629
FT                   /*tag= c
FT                   /note= "C changed to A to create a ScaI site"
XX
PN   WO9838217-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;    98WO-US03991.
XX
PR   07-OCT-1997;    97US-0061268.
PR   27-FEB-1997;    97US-0039465.
XX
PA   (EGAN/) EGAN C C.
PA   (HERR/) HERRICK-DAVIS K.
PA   (TEIT/) TEITLER M.
XX
PI   Egan CC,  Herrick-Davis K,  Teitler M;
XX
DR   WPI; 1998-495389/42.
XX
PT   Method of constitutively activating targeted G-protein coupled
PT   mono:amine receptor - comprises use of site directed mutagenesis,
PT   useful for, e.g. screening for agonists and antagonists of native
PT   receptor
XX
PS   Claim 19; Fig 33; 97pp; English.
XX
CC   The present sequence encodes a rat 5-HT2C serotonin receptor S312K
CC   mutant. The present sequence contains a ScaI restriction site. 5-HT2C
CC   is a G-protein-coupled receptor. The sequence of the wild type receptor
CC   (AAV48279) is modified using the method of the invention to produce the
CC   present sequence. The mutant receptor exists in a constitutively
CC   activated state exhibiting both a greater response to agonists and a
CC   coupling to the G protein second messenger system even in the absence
CC   of agonists. The constitutively activated monoamine G protein-coupled
CC   receptor can be used to screen for agonists, inverse agonists, and
CC   antagonists of the native receptor.
XX
SQ   Sequence 2246 BP; 540 A; 540 C; 539 G; 627 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 19;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 15
AAV48279/c
ID   AAV48279 standard; DNA; 2246 BP.
XX
AC   AAV48279;
XX
DT   21-JAN-1999  (first entry)
XX
DE   Rat 5-HT2C serotonin receptor DNA.
XX
KW   Rat 5-HT2C serotonin receptor; G-protein-coupled receptor;
KW   constitutively activated monoamine G protein-coupled receptor;
KW   screen; agonist; inverse agonist; antagonist; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             688..2070
FT                   /*tag= a
FT                   /product= 5-HT2C serotonin receptor
XX
PN   WO9838217-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;    98WO-US03991.
XX
PR   07-OCT-1997;    97US-0061268.
PR   27-FEB-1997;    97US-0039465.
XX
PA   (EGAN/) EGAN C C.
PA   (HERR/) HERRICK-DAVIS K.
PA   (TEIT/) TEITLER M.
XX
PI   Egan CC,  Herrick-Davis K,  Teitler M;
XX
DR   WPI; 1998-495389/42.
DR   P-PSDB; AAW77105.
XX
PT   Method of constitutively activating targeted G-protein coupled
PT   mono:amine receptor - comprises use of site directed mutagenesis,
PT   useful for, e.g. screening for agonists and antagonists of native
PT   receptor
XX
PS   Disclosure; Fig 2A; 97pp; English.
XX
CC   The present sequence encodes a rat 5-HT2C serotonin receptor. This
CC   is a G-protein-coupled receptor. The sequence of the receptor is
CC   modified (see AAV48288-91) using the method of the invention so
CC   that the receptor exists in a constitutively activated state exhibiting
CC   both a greater response to agonists and a coupling to the G protein
CC   second messenger system even in the absence of agonists. The
CC   constitutively activated monoamine G protein-coupled receptor can be
CC   used to screen for agonists, inverse agonists, and antagonists of the
CC   native receptor.
XX
SQ   Sequence 2246 BP; 537 A; 543 C; 538 G; 628 T; 0 other;

  Query Match             83.2%;  Score 15.8;  DB 19;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


Search completed: January 21, 2003, 16:28:00
Job time : 161.389 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:52:32 ; Search time 1198.97 Seconds
(without alignments)
229.633 Million cell updates/sec

Title: US-09-853-688-35
Perfect score: 17
Sequence: 1 ggqagccccagcaatgc 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	16	94.1	461	10	AW558892	AW558892 L0301D12-
	2	16	94.1	619	17	BH108836	BH108836 RPCI-24-2
c	3	16	94.1	621	12	BG324158	BG324158 602423116
c	4	16	94.1	636	10	AW239325	AW239325 xb39a04.y
	5	16	94.1	642	12	BG076347	BG076347 H3158F07-
	6	16	94.1	774	12	BG704650	BG704650 602688222
c	7	16	94.1	976	13	BM468789	BM468789 AGENCOURT
	8	15.4	90.6	300	9	AA852498	AA852498 NHTBCae14
c	9	15.4	90.6	319	10	AW067967	AW067967 cn22h11.x
c	10	15.4	90.6	320	9	AI449767	AI449767 mr69a08.x
c	11	15.4	90.6	336	9	AA629118	AA629118 af57e02.s
c	12	15.4	90.6	382	17	BH842232	BH842232 TC3-50P20
c	13	15.4	90.6	382	17	BH845093	BH845093 TC3-50P20
c	14	15.4	90.6	383	13	BI005357	BI005357 CM0-HN020
c	15	15.4	90.6	385	17	BI14181	BI14181 A-1054B4.TP
c	16	15.4	90.6	393	9	AI752118	AI752118 cn13g09.y
	17	15.4	90.6	406	12	BF406319	BF406319 UI-R-CA1-
	18	15.4	90.6	411	9	AA536528	AA536528 LD17175.5
c	19	15.4	90.6	411	12	BF856554	BF856554 PM3-FT002
	20	15.4	90.6	412	14	R45494	R45494 yg42a09.s1
c	21	15.4	90.6	413	9	AI751985	AI751985 cn13a05.x
c	22	15.4	90.6	413	12	BF857617	BF857617 PM3-FT020
	23	15.4	90.6	416	13	BI239628	BI239628 RE36172.5
c	24	15.4	90.6	420	9	AI751499	AI751499 cn10c11.y
c	25	15.4	90.6	424	9	AI751854	AI751854 cn12c05.x
c	26	15.4	90.6	425	10	BB268875	BB268875 BB268875
c	27	15.4	90.6	440	9	AI751606	AI751606 cn10h08.y
	28	15.4	90.6	454	9	AA391288	AA391288 LD09888.5
c	29	15.4	90.6	457	9	AI750858	AI750858 cn06a12.x
	30	15.4	90.6	468	13	BI243443	BI243443 RE41086.5
	31	15.4	90.6	501	14	BM727488	BM727488 UI-E-EJ0-
	32	15.4	90.6	512	13	BI363494	BI363494 RE48317.5
	33	15.4	90.6	512	13	BI403236	BI403236 MI-P-CP1-
	34	15.4	90.6	537	13	BI371847	BI371847 RE58854.5
c	35	15.4	90.6	540	9	AU144696	AU144696 AU144696
	36	15.4	90.6	549	13	BI238029	BI238029 RE34390.5
	37	15.4	90.6	562	13	BI486096	BI486096 RE69469.5
	38	15.4	90.6	581	14	W22315	W22315 65D2 Human
	39	15.4	90.6	584	9	AA160661	AA160661 zq49h09.r
	40	15.4	90.6	587	9	AI520081	AI520081 LD40038.5
c	41	15.4	90.6	589	12	BG162438	BG162438 de34e08.y
	42	15.4	90.6	591	14	BM931980	BM931980 UI-E-EJ1-
c	43	15.4	90.6	594	14	BM685020	BM685020 UI-E-EJ1-
	44	15.4	90.6	597	13	BI239575	BI239575 RE36110.5
	45	15.4	90.6	607	13	BI162042	BI162042 RE01467.5

ALIGNMENTS

RESULT 1
AW558892
LOCUS AW558892 461 bp mRNA linear EST 31-AUG-2000
DEFINITION L0301D12-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA clone L0301D12 3', mRNA sequence.
ACCESSION AW558892
VERSION AW558892.1 GI:7204321
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 461)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0301 row: D column: 12
Seq primer: -21M13 Forward
High quality sequence stop: 461
POLYA=Yes.

```
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:L0301D12-3"
                     /db_xref="taxon:10090"
                     /clone="L0301D12"
                     /clone_lib="NIA Mouse Newborn Ovary cDNA Library"
                     /sex="female"
                     /dev_stage="Newborn Ovary"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                     Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                     7 Newborn Ovary. The double-stranded cDNA was synthesized
                     by Gibco's kit with an Oligo(dT) primer [NotI
                     primer-adapter from GibcoBRL]
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from
                     2.56ug of total RNA . The double-stranded cDNAs were
                     treated with T4 DNA polymerase and purified by
                     ethanol-precipitation. The cDNAs were ligated to
                     Lone-linker LL-Sal3 (include SalI sequence). The cDNAs
                     were purified by phenol/chloroform and separated from
                     free linkers by Centricon 100. Then, cDNAs were amplified
                     by long-range high fidelity PCR using Takara's Ex Taq
                     polymerase. Then, the cDNAs were purified by
                     phenol/chloroform and by Centricon 100. The cDNAs were
                     digested with SalI and NotI enzymes. Then, the cDNAs were
                     size selected by Gibco's Size Fractionation Column. The
                     cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by chemical method. The library was
                     constructed by Xiaohong Wang and Yulan Piao."
BASE COUNT       125 a     86 c    126 g    124 t
ORIGIN

  Query Match             94.1%;  Score 16;  DB 10;  Length 461;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATG 16
           ||||||||||||||||
Db     280 GGGAGCCCCAGCAATG 295


RESULT 2
BH108836
LOCUS       BH108836                619 bp    DNA     linear   GSS 19-JUL-2001
DEFINITION  RPCI-24-238C2.TV RPCI-24 Mus musculus genomic clone RPCI-24-238C2,
            DNA sequence.
ACCESSION   BH108836
VERSION     BH108836.1  GI:14941302
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-238C2.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 238  row: C  column: 2
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .619
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-238C2"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
BASE COUNT       138 a    149 c    172 g    160 t
ORIGIN

  Query Match             94.1%;  Score 16;  DB 17;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATG 16
           ||||||||||||||||
Db     207 GGGAGCCCCAGCAATG 222


RESULT 3
BG324158/c
LOCUS       BG324158                621 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602423116F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4561371 5',
            mRNA sequence.
ACCESSION   BG324158
VERSION     BG324158.1  GI:13130595
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1271  row: o  column: 04
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4561371"
                     /clone_lib="NIH_MGC_14"
                     /tissue_type="renal cell adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
```

(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 128 a 185 c 188 g 120 t
ORIGIN

Query Match 94.1%; Score 16; DB 12; Length 621;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGAGCCCCAGCAATGC 17
||||||||||||||||
Db 41 GGAGCCCCAGCAATGC 16

RESULT 4
AW239325/c
LOCUS AW239325 636 bp mRNA linear EST 13-DEC-1999
DEFINITION xb39a04.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2578638 5' similar to gb:J03934 NAD(P)H DEHYDROGENASE (HUMAN);, mRNA sequence.
ACCESSION AW239325
VERSION AW239325.1 GI:6571715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: xb39a04.x1
Contact: Robert Strausberg, Ph.D.
Email: cgaphs-r@mail.nih.gov
Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40RP from Gibco
High quality sequence stop: 437.
FEATURES Location/Qualifiers
source 1. .636
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2578638"
/clone_lib="NCI_CGAP_Lu31"
/sex="male"
/dev_stage="fetal, 14 wk post-conception"
/lab_host="DH10B"
/note="Organ: lung, cell line; Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally, no 5' adaptor. Primer: Oligo dT. Full-length library constructed by Life Technologies."
BASE COUNT 154 a 170 c 168 g 143 t 1 others
ORIGIN

Query Match 94.1%; Score 16; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGAGCCCCAGCAATGC 17
||||||||||||||||
Db 635 GGAGCCCCAGCAATGC 620

RESULT 5
BG076347
LOCUS BG076347 642 bp mRNA linear EST 26-JAN-2001
DEFINITION H3158F07-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3158F07 3', mRNA sequence.
ACCESSION BG076347
VERSION BG076347.1 GI:12558916
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 642)
AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka,T.S., Carter,M.G. and Ko,M.S.H.
TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other_ESTs: H3158F07-5
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3158 row: F column: 07
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3158F07-3"
/db_xref="taxon:10090"
/clone="H3158F07"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse develolpment, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
BASE COUNT 166 a 135 c 181 g 160 t
ORIGIN

Query Match 94.1%; Score 16; DB 12; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGCCCCAGCAATG 16
||||||||||||||||
Db 280 GGGAGCCCCAGCAATG 295

RESULT 6
BG704650/c
LOCUS BG704650 774 bp mRNA linear EST 07-MAY-2001
DEFINITION 602688222F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4820860 5', mRNA sequence.
ACCESSION BG704650
VERSION BG704650.1 GI:13978201

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10726  row: k  column: 05
            High quality sequence stop: 700.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4820860"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      171 a    216 c    233 g    154 t
ORIGIN

  Query Match             94.1%;  Score 16;  DB 12;  Length 774;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           ||||||||||||||||
Db      54 GGAGCCCCAGCAATGC 39


RESULT 7
BM468789/c
LOCUS       BM468789                 976 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6481239 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5541510
            5', mRNA sequence.
ACCESSION   BM468789
VERSION     BM468789.1  GI:18517831
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 976)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12239  row: f  column: 07
            High quality sequence stop: 642.
FEATURES             Location/Qualifiers
     source          1. .976
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5541510"
                     /clone_lib="NIH_MGC_88"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      202 a    289 c    267 g    218 t
ORIGIN

  Query Match             94.1%;  Score 16;  DB 13;  Length 976;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATG 16
           ||||||||||||||||
Db     705 GGGAGCCCCAGCAATG 690


RESULT 8
AA852498
LOCUS       AA852498                 300 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  NHTBCael4a07f1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBCael4a07, mRNA sequence.
ACCESSION   AA852498
VERSION     AA852498.1  GI:2941091
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
            ,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NHTBCael4a07"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
BASE COUNT       64 a    102 c     99 g     34 t      1 others
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 9;  Length 300;
  Best Local Similarity   94.1%;  Pred. No. 3.5e+03;
```

```
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db     167 GGGAGCCCCAGCAACGC 183


RESULT 9
AW067967/c
LOCUS       AW067967                319 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  cn22h11.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn22h11 random, mRNA sequence.
ACCESSION   AW067967
VERSION     AW067967.1  GI:6022965
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 319)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
            ,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 22  row: h  column: 11
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .319
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn22h11"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
BASE COUNT       41 a     96 c    112 g     70 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 319;
  Best Local Similarity   94.1%;  Pred. No. 3.6e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db     178 GGGAGCCCCAGCAACGC 162


RESULT 10
AI449767/c
LOCUS       AI449767                320 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  mr69a08.x1 Stratagene mouse testis (#937308) Mus musculus cDNA
            clone IMAGE:602678 3', mRNA sequence.
ACCESSION   AI449767
VERSION     AI449767.1  GI:4292306
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Possible reversed clone: polyT not found.
FEATURES             Location/Qualifiers
     source          1. .320
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:602678"
                     /clone_lib="Stratagene mouse testis (#937308)"
                     /sex="males"
                     /tissue_type="testis"
                     /dev_stage="10-12 week old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
                     -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       81 a     70 c     85 g     84 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 9;  Length 320;
  Best Local Similarity   94.1%;  Pred. No. 3.6e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||| ||||||
Db      34 GGGAGCCCCACCAATGC 18


RESULT 11
AA629118/c
LOCUS       AA629118                336 bp    mRNA    linear   EST 02-MAR-1998
DEFINITION  af57e02.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:1035770 3' similar to gb:M23102 HIGH AFFINITY NERVE GROWTH
            FACTOR RECEPTOR PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AA629118
VERSION     AA629118.1  GI:2541505
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
```

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 710    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1035770"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       64 a     95 c     98 g     79 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 9;  Length 336;
  Best Local Similarity   94.1%;  Pred. No. 3.6e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db     301 GGGAGCCCCAGCAACGC 285


RESULT 12
BH842232/c
LOCUS       BH842232                382 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  TC3-50P20.TP TC3 Trypanosoma cruzi genomic clone TC3-50P20, DNA
            sequence.
ACCESSION   BH842232
VERSION     BH842232.1  GI:21409447
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi.
  ORGANISM  Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Myler,P.J., Aggarwal,G., Fazelinia,G., Mack,J., Marty,A., Munden,H.
            , Nelson,S., Pentony,M., Rinta,J., Robertson,L., Seyler,A., Sisk,E.
            , Stuart,K., Vogt,C., Worthey,E., El-Sayed,N.M., Ghedin,E. and
            Andersson,B.
  TITLE     Trypanosoma cruzi CL-Brener TC3 BAC-end sequencing
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: TC3-50P20.TP.1 TC3-50P20.TV
            Contact: Peter Myler
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seattle, WA 98109, USA
            Tel: 206 284 8846
            Fax: 206 284 0313
            Email: mylerpj@sbri.org
            Clones are derived from the Trypanosoma cruzi CL-Brener BAC library
            TC3. For clone availability, please contact Dr. Bjorn Andersson at
            Uppsala University (bjorn.andersson@genpat.uu.se).
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Trypanosoma cruzi"
                     /strain="CL Brener"
                     /db_xref="taxon:5693"
                     /clone="TC3-50P20"
                     /clone_lib="TC3"
                     /note="Vector: pBeloBAC11; Site_1: Hin dIII; Constructed
                     for Uppsala University by Marie-Christine Le Paslier in
                     the laboratory of Denis Le Paslier at the Centre d'Etude
                     du Polymorphisme Humain (CEPH), Paris, France. Briefly,
                     Trypanosoma cruzi CL-Brener agarose embedded DNA (obtained
                     from Dr. Franco da Silveira) was partially digested with
                     Hin dIII.  High molecular weight fragments were ligated in
                     pBeloBAC11 digested with Hin dIII.  The average insert
                     size is 100 kb. Total clone coverage: approx. 33 X the
                     haploid genome."
BASE COUNT       88 a     112 c     81 g     101 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 17;  Length 382;
  Best Local Similarity   94.1%;  Pred. No. 3.8e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||| |||||||
Db     172 GGGAGCCCCTGCAATGC 156


RESULT 13
BH845093/c
LOCUS       BH845093                382 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  TC3-50P20.TP.1 TC3 Trypanosoma cruzi genomic clone TC3-50P20, DNA
            sequence.
ACCESSION   BH845093
VERSION     BH845093.1  GI:21415281
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi.
  ORGANISM  Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Myler,P.J., Aggarwal,G., Fazelinia,G., Mack,J., Marty,A., Munden,H.
            , Nelson,S., Pentony,M., Rinta,J., Robertson,L., Seyler,A., Sisk,E.
            , Stuart,K., Vogt,C., Worthey,E., El-Sayed,N.M., Ghedin,E. and
            Andersson,B.
  TITLE     Trypanosoma cruzi CL-Brener TC3 BAC-end sequencing
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: TC3-50P20.TP TC3-50P20.TV
            Contact: Peter Myler
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seattle, WA 98109, USA
            Tel: 206 284 8846
            Fax: 206 284 0313
            Email: mylerpj@sbri.org
            Clones are derived from the Trypanosoma cruzi CL-Brener BAC library
            TC3. For clone availability, please contact Dr. Bjorn Andersson at
            Uppsala University (bjorn.andersson@genpat.uu.se).
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Trypanosoma cruzi"
                     /strain="CL Brener"
                     /db_xref="taxon:5693"
                     /clone="TC3-50P20"
                     /clone_lib="TC3"
                     /note="Vector: pBeloBAC11; Site_1: Hin dIII; Constructed
                     for Uppsala University by Marie-Christine Le Paslier in
                     the laboratory of Denis Le Paslier at the Centre d'Etude
                     du Polymorphisme Humain (CEPH), Paris, France. Briefly,
                     Trypanosoma cruzi CL-Brener agarose embedded DNA (obtained
                     from Dr. Franco da Silveira) was partially digested with
                     Hin dIII.  High molecular weight fragments were ligated in
                     pBeloBAC11 digested with Hin dIII.  The average insert
                     size is 100 kb. Total clone coverage: approx. 33 X the
                     haploid genome."
BASE COUNT       88 a     112 c     81 g     101 t
ORIGIN
```

```
  Query Match             90.6%;  Score 15.4;  DB 17;  Length 382;
  Best Local Similarity   94.1%;  Pred. No. 3.8e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||| |||||||
Db     172 GGGAGCCCCTGCAATGC 156


RESULT 14
BI005357/c
LOCUS       BI005357                 383 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  CM0-HN0206-090401-813-e11 HN0206 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI005357
VERSION     BI005357.1  GI:14409431
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM0&t2=CM0-HN0206-
            090401-813-e11&t3=2001-04-09&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 383.
FEATURES             Location/Qualifiers
     source          1. .383
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HN0206"
                     /dev_stage="Adult"
                     /note="Organ: head_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       74 a     97 c     89 g    123 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 13;  Length 383;
  Best Local Similarity   94.1%;  Pred. No. 3.8e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           || ||||||||||||||
Db      54 GGCAGCCCCAGCAATGC 38


RESULT 15
BI418l/c
LOCUS       BI418l                   385 bp    DNA     linear   GSS 18-JUN-1998
DEFINITION  A-1054B4.TP CIT978SK Homo sapiens genomic clone A-1054B4, DNA
            sequence.
ACCESSION   BI418l
VERSION     BI418l.1  GI:2121930
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: 1054B4.TV
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="A-1054B4"
                     /clone_lib="CIT978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
BASE COUNT       80 a     126 c     94 g     85 t
ORIGIN

  Query Match             90.6%;  Score 15.4;  DB 17;  Length 385;
  Best Local Similarity   94.1%;  Pred. No. 3.8e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| |
Db     154 GGGAGCCCCAGCAATTC 138


Search completed: January 21, 2003, 15:38:03
Job time : 1203.97 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:55:32 ; Search time 30.6944 Seconds
(without alignments)
246.942 Million cell updates/sec

Title: US-09-853-688-35
Perfect score: 17
Sequence: 1 gggagccccagcaatgc 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17	100.0	17	10	US-09-853-688-35	Sequence 35, Appl
	2	17	100.0	3700	10	US-09-853-688-5	Sequence 5, Appli
	3	15.4	90.6	468	10	US-09-864-761-5528	Sequence 5528, Ap
	4	15.4	90.6	2820	10	US-09-924-859A-4	Sequence 4, Appli
c	5	15.4	90.6	5086	10	US-09-880-107-3947	Sequence 3947, Ap
c	6	15.4	90.6	5145	10	US-09-925-299-206	Sequence 206, App
c	7	15.4	90.6	5416	10	US-09-954-456-786	Sequence 786, App
c	8	15.4	90.6	5416	10	US-09-880-107-2094	Sequence 2094, Ap
c	9	15.4	90.6	5432	12	US-10-044-090-22	Sequence 22, Appl
	10	15	88.2	1041	9	US-09-738-626-3422	Sequence 3422, Ap
c	11	15	88.2	1152	9	US-09-738-626-3424	Sequence 3424, Ap
	12	15	88.2	1894	10	US-09-126-945B-1	Sequence 1, Appli
	13	15	88.2	1894	10	US-09-841-963A-1	Sequence 1, Appli
	14	15	88.2	1905	10	US-09-866-356-2	Sequence 2, Appli
	15	14.4	84.7	266	10	US-09-962-436-91	Sequence 91, Appl
	16	14.4	84.7	302	9	US-10-015-219-718	Sequence 718, App
	17	14.4	84.7	302	10	US-09-777-564-718	Sequence 718, App
	18	14.4	84.7	316	10	US-09-998-598-1147	Sequence 1147, Ap
	19	14.4	84.7	698	10	US-09-910-943-311	Sequence 311, App
	20	14.4	84.7	1570	9	US-09-798-051-4	Sequence 4, Appli
	21	14.4	84.7	1720	9	US-09-965-528-39	Sequence 39, Appl
	22	14.4	84.7	1732	9	US-10-063-547-81	Sequence 81, Appl
	23	14.4	84.7	1732	9	US-10-174-590-281	Sequence 281, App
	24	14.4	84.7	1732	9	US-10-176-758-281	Sequence 281, App
	25	14.4	84.7	1732	12	US-10-006-867-81	Sequence 81, Appl
	26	14.4	84.7	1732	12	US-10-052-586-281	Sequence 281, App
	27	14.4	84.7	1839	9	US-09-798-051-1	Sequence 1, Appli
	28	14	82.4	1734	9	US-10-103-511-3	Sequence 3, Appli
	29	14	82.4	1734	10	US-09-805-204-3	Sequence 3, Appli
c	30	14	82.4	4065	10	US-09-808-571A-3	Sequence 3, Appli
c	31	14	82.4	4341	10	US-09-808-571A-1	Sequence 1, Appli
c	32	14	82.4	9745	10	US-09-764-869-2259	Sequence 2259, Ap
c	33	14	82.4	12149	10	US-09-764-869-2258	Sequence 2258, Ap
c	34	14	82.4	29629	12	US-10-135-689-3	Sequence 3, Appli
c	35	14	82.4	170834	10	US-09-835-232-7	Sequence 7, Appli
c	36	13.8	81.2	293	10	US-09-983-965-2249	Sequence 2249, Ap
c	37	13.8	81.2	337	10	US-09-983-965-2463	Sequence 2463, Ap
c	38	13.8	81.2	349	10	US-09-983-965-2493	Sequence 2493, Ap
c	39	13.8	81.2	362	10	US-09-983-965-2530	Sequence 2530, Ap
c	40	13.8	81.2	365	9	US-09-796-692-5576	Sequence 5576, Ap
c	41	13.8	81.2	379	10	US-09-783-590-7717	Sequence 7717, Ap
c	42	13.8	81.2	384	9	US-10-040-739-1302	Sequence 1302, Ap
	43	13.8	81.2	390	10	US-09-960-352-3031	Sequence 3031, Ap
c	44	13.8	81.2	398	10	US-09-728-445-656	Sequence 656, App
c	45	13.8	81.2	479	9	US-09-796-692-3578	Sequence 3578, Ap

ALIGNMENTS

RESULT 1
US-09-853-688-35
; Sequence 35, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; TITLE OF INVENTION: HUMANS, THE VARIATIONS AND THEIR USES
; FILE REFERENCE: WCM78
; CURRENT APPLICATION NUMBER: US/09/853,688
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-688-35

Query Match 100.0%; Score 17; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGCCCCAGCAATGC 17
||||||||||||||||||
Db 1 GGGAGCCCCAGCAATGC 17

RESULT 2
US-09-853-688-5
; Sequence 5, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; TITLE OF INVENTION: HUMANS, THE VARIATIONS AND THEIR USES

```
;  FILE REFERENCE: WCM78
;  CURRENT APPLICATION NUMBER: US/09/853,688
;  CURRENT FILING DATE:  2001-05-14
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 3700
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-853-688-5

  Query Match             100.0%;  Score 17;  DB 10;  Length 3700;
  Best Local Similarity   100.0%;  Pred. No. 7.8;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||||
Db      86 GGGAGCCCCAGCAATGC 102


RESULT 3
US-09-864-761-5528
; Sequence 5528, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5528
;   LENGTH: 468
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005529.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
US-09-864-761-5528

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 43;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 4
US-09-924-859A-4
; Sequence 4, Application US/09924859A
; Patent No. US20020137113A1
; GENERAL INFORMATION:
;  APPLICANT: Godowski, Paul J.
;  APPLICANT:  Mark, Melanie R.
;  APPLICANT:  Sadick, Michael D.
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Wong, Wai Lee Tan
;  TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
;  FILE REFERENCE: P0854C1P2C1
;  CURRENT APPLICATION NUMBER: US/09/924,859A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US/09/417,381
;  PRIOR FILING DATE: 1999-10-13
;  NUMBER OF SEQ ID NOS: 11
; SEQ ID NO 4
;   LENGTH: 2820
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-924-859A-4

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 2820;
  Best Local Similarity   94.1%;  Pred. No. 50;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||| ||
Db    2643 GGGAGCCCCAGCAACGC 2659


RESULT 5
US-09-880-107-3947/c
; Sequence 3947, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
```

```
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3947
;    LENGTH: 5086
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z74616
;    NAME/KEY: unsure
;    LOCATION: (1)..(5086)
;    OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3947

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 5086;
  Best Local Similarity   94.1%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGAGCCCCAGCAATGC 17
          ||||||||||||||| ||
Db   1102 GGGAGCCCCAGCAACGC 1086


RESULT 6
US-09-925-299-206/c
; Sequence 206, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
;    LENGTH: 5145
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (4)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (17)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (5126)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (5143)
;    OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-206

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 5145;
  Best Local Similarity   94.1%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGAGCCCCAGCAATGC 17
          ||||||||||||||| ||
Db   1105 GGGAGCCCCAGCAACGC 1089


RESULT 7
US-09-954-456-786/c
; Sequence 786, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 786
;    LENGTH: 5416
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-954-456-786

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 5416;
  Best Local Similarity   94.1%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGAGCCCCAGCAATGC 17
          ||||||||||||||| ||
Db   1432 GGGAGCCCCAGCAACGC 1416


RESULT 8
US-09-880-107-2094/c
; Sequence 2094, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2094
;    LENGTH: 5416
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: Genbank Accession No. US20020142981A1 J03464
US-09-880-107-2094

  Query Match             90.6%;  Score 15.4;  DB 10;  Length 5416;
  Best Local Similarity   94.1%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGAGCCCCAGCAATGC 17
```

```
              |||||||||||||||  ||
Db    1412 GGGAGCCCCAGCAACGC 1416


RESULT 9
US-10-044-090-22/c
; Sequence 22, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 22
;   LENGTH: 5432
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1383093.13
US-10-044-090-22

  Query Match             90.6%;  Score 15.4;  DB 12;  Length 5432;
  Best Local Similarity   94.1%;  Pred. No. 52;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||  ||
Db    1435 GGGAGCCCCAGCAACGC 1419


RESULT 10
US-09-738-626-3422
; Sequence 3422, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3422
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-3422

  Query Match             88.2%;  Score 15;  DB 9;  Length 1041;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 GAGCCCCAGCAATGC 17
           |||||||||||||||
Db     496 GAGCCCCAGCAATGC 510


RESULT 11
US-09-738-626-3424/c
; Sequence 3424, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3424
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-3424

  Query Match             88.2%;  Score 15;  DB 9;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 GAGCCCCAGCAATGC 17
           |||||||||||||||
Db     858 GAGCCCCAGCAATGC 844


RESULT 12
US-09-126-945B-1
; Sequence 1, Application US/09126945B
; Patent No. US20010010934A1
; GENERAL INFORMATION:
;  APPLICANT:  Libermann, Towia A.
;  APPLICANT:  Oettgen, Joerg P.
;  APPLICANT:  Kunsch, Charles A.
;  APPLICANT:  Endress, Gregory A.
;  APPLICANT:  Rosen, Craig A.
;  TITLE OF INVENTION:  Prostate Derived Ets Factor
;  FILE REFERENCE:  1488.1090000
;  CURRENT APPLICATION NUMBER: US/09/126,945B
;  CURRENT FILING DATE:  1998-07-31
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-126-945B-1

  Query Match             88.2%;  Score 15;  DB 10;  Length 1894;
  Best Local Similarity   100.0%;  Pred. No. 77;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATG 16
           |||||||||||||||
```

```
Db    864 GGAGCCCCAGCAATG 878


RESULT 13
US-09-841-963A-1
; Sequence 1, Application US/09841963A
; Patent No. US20020081601A1
; GENERAL INFORMATION:
;  APPLICANT: Watson, Dennis K.
;  APPLICANT:  Papas, Takis S. (Deceased)
;  APPLICANT:  Papas, Tula C. (Legal Representative)
;  TITLE OF INVENTION: Methods and compositions for the diagnosis and treatment of cance
;  TITLE OF INVENTION:  based on transcription factor ETS2
;  FILE REFERENCE: 10545-015-999
;  CURRENT APPLICATION NUMBER: US/09/841,963A
;  CURRENT FILING DATE:  2001-04-25
;  PRIOR APPLICATION NUMBER: PCT/US99/27805
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: 06/109,850
;  PRIOR FILING DATE: 1998-11-25
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (416)..(1423)
US-09-841-963A-1

  Query Match             88.2%;  Score 15;  DB 10;  Length 1894;
  Best Local Similarity   100.0%;  Pred. No. 77;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GGAGCCCCAGCAATG 16
          |||||||||||||||
Db    864 GGAGCCCCAGCAATG 878


RESULT 14
US-09-866-356-2
; Sequence 2, Application US/09866356
; Patent No. US20020098543A1
;   GENERAL INFORMATION:
;        APPLICANT: Bandman, Olga
;                   Corley, Neil C.
;                   Guegler, Karl J.
;                   Lal, Preeti
;        TITLE OF INVENTION: PROSTATE-ASSOCIATED ETS PROTEIN
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/866,356
;             FILING DATE: 29-May-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/055,113
;             FILING DATE: 1998-04-03
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0501 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1905 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: PROSTUT12
;             CLONE: 1813005
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-866-356-2

  Query Match             88.2%;  Score 15;  DB 10;  Length 1905;
  Best Local Similarity   100.0%;  Pred. No. 77;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GGAGCCCCAGCAATG 16
          |||||||||||||||
Db    878 GGAGCCCCAGCAATG 892


RESULT 15
US-09-962-436-91
; Sequence 91, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
;  APPLICANT: Soppet, Daniel
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using S
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-75
;  CURRENT APPLICATION NUMBER: US/09/962,436
;  CURRENT FILING DATE:  2001-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,082
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/234,924
;  PRIOR FILING DATE: 2000-09-25
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: n=a,t,g or c
US-09-962-436-91

  Query Match             84.7%;  Score 14.4;  DB 10;  Length 266;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GGAGCCCCAGCAATGC 17
          ||||| ||||||||||
Db    170 GGAGCTCCAGCAATGC 185


Search completed: January 21, 2003, 16:29:15
Job time : 30.6944 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 15:09:37 ; Search time 29.2778 Seconds
(without alignments)
178.070 Million cell updates/sec

Title: US-09-853-688-35
Perfect score: 17
Sequence: 1 gggagccccagcaatgc 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.4	90.6	2820	1	US-08-286-305A-4	Sequence 4, Appli
	2	15.4	90.6	2820	2	US-08-441-104A-4	Sequence 4, Appli
	3	15.4	90.6	2820	2	US-08-440-816A-4	Sequence 4, Appli
	4	15.4	90.6	2820	4	US-09-417-381A-4	Sequence 4, Appli
c	5	15.4	90.6	38682	4	US-08-943-731-2	Sequence 2, Appli
	6	15	88.2	1905	4	US-09-055-113-2	Sequence 2, Appli
c	7	15	88.2	1929	2	US-09-016-000-10	Sequence 10, Appl
	8	14.4	84.7	23	2	US-08-468-551-9	Sequence 9, Appli
c	9	14.4	84.7	104	3	US-08-717-294-104	Sequence 104, App
	10	14.4	84.7	4451	3	US-08-717-294-42	Sequence 42, Appl
	11	14.4	84.7	9179	4	US-09-453-702B-162	Sequence 162, App
c	12	14.4	84.7	12537	2	US-08-611-280-4	Sequence 4, Appli
c	13	14.4	84.7	12537	4	US-09-195-940-4	Sequence 4, Appli
c	14	14.4	84.7	12537	4	US-09-562-466-4	Sequence 4, Appli
	15	14	82.4	1209	4	US-08-822-774-18	Sequence 18, Appl
	16	14	82.4	1209	4	US-09-632-711-18	Sequence 18, Appl
	17	14	82.4	1209	4	US-09-632-703B-18	Sequence 18, Appl
	18	14	82.4	1209	4	US-09-632-702-18	Sequence 18, Appl
	19	14	82.4	7379	4	US-09-341-587-5	Sequence 5, Appli
c	20	14	82.4	29629	4	US-09-729-995-3	Sequence 3, Appli
c	21	13.8	81.2	766	4	US-09-105-839D-1	Sequence 1, Appli
	22	13.8	81.2	2062	1	US-08-050-319B-24	Sequence 24, Appl
	23	13.8	81.2	2062	2	US-08-465-982-24	Sequence 24, Appl
	24	13.8	81.2	2081	2	US-09-096-982-7	Sequence 7, Appli
	25	13.8	81.2	2081	2	US-08-653-650A-7	Sequence 7, Appli
	26	13.8	81.2	2161	3	US-09-106-038A-1	Sequence 1, Appli
	27	13.8	81.2	2161	4	US-09-505-250-3	Sequence 3, Appli
	28	13.8	81.2	2175	1	US-08-321-668-1	Sequence 1, Appli
	29	13.8	81.2	2175	1	US-08-837-941-1	Sequence 1, Appli
	30	13.8	81.2	2175	1	US-08-126-016-1	Sequence 1, Appli
	31	13.8	81.2	2175	4	US-08-054-970-1	Sequence 1, Appli
	32	13.8	81.2	2230	4	US-09-189-527-12	Sequence 12, Appl
	33	13.8	81.2	2834	4	US-09-305-384-6	Sequence 6, Appli
	34	13.8	81.2	3196	2	US-09-096-982-4	Sequence 4, Appli
	35	13.8	81.2	3196	2	US-08-653-650A-4	Sequence 4, Appli
	36	13.8	81.2	6235	4	US-09-305-384-5	Sequence 5, Appli
	37	13.8	81.2	6679	4	US-09-305-384-1	Sequence 1, Appli
c	38	13.4	78.8	189	1	US-08-473-981A-1	Sequence 1, Appli
c	39	13.4	78.8	189	2	US-08-474-087-1	Sequence 1, Appli
	40	13.4	78.8	680	4	US-08-943-731-92	Sequence 92, Appl
c	41	13.4	78.8	777	4	US-08-998-416-247	Sequence 247, App
	42	13.4	78.8	1491	4	US-09-662-249A-3	Sequence 3, Appli
c	43	13.4	78.8	1781	1	US-08-314-615-2	Sequence 2, Appli
c	44	13.4	78.8	1781	1	US-08-314-362-2	Sequence 2, Appli
c	45	13.4	78.8	1781	1	US-08-433-010-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-286-305A-4
; Sequence 4, Application US/08286305A
; Patent No. 5766863
;  GENERAL INFORMATION:
;    APPLICANT:  Godowski, Paul J.
;    APPLICANT:  Mark, Melanie R.
;    APPLICANT:  Sadick, Michael D.
;    APPLICANT:  Shelton, David L.
;    APPLICANT:  Wong, Wai Lee Tan
;    TITLE OF INVENTION:  KINASE RECEPTOR ACTIVATION ASSAY
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,305A
;      FILING DATE:  05-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/170558
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/157563
;      FILING DATE:  23-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  854C1P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2820 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-286-305A-4
```

```
  Query Match             90.6%;  Score 15.4;  DB 1;  Length 2820;
  Best Local Similarity   94.1%;  Pred. No. 40;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db    2643 GGGAGCCCCAGCAACGC 2659


RESULT 2
US-08-441-104A-4
; Sequence 4, Application US/08441104A
; Patent No. 5891650
;  GENERAL INFORMATION:
;    APPLICANT:  Godowski, Paul J.
;    APPLICANT:  Mark, Melanie R.
;    APPLICANT:  Sadick, Michael D.
;    APPLICANT:  Shelton, David L.
;    APPLICANT:  Wong, Wai Lee Tan
;    TITLE OF INVENTION:  KINASE RECEPTOR ACTIVATION ASSAY
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 720 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,104A
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286305
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/170558
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/157563
;      FILING DATE:  23-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  854C1P1C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2820 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-441-104A-4

  Query Match             90.6%;  Score 15.4;  DB 2;  Length 2820;
  Best Local Similarity   94.1%;  Pred. No. 40;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db    2643 GGGAGCCCCAGCAACGC 2659


RESULT 3
US-08-440-816A-4
; Sequence 4, Application US/08440816A
; Patent No. 5914237
;  GENERAL INFORMATION:
;    APPLICANT:  Godowski, Paul J.
;    APPLICANT:  Mark, Melanie R.
;    APPLICANT:  Sadick, Michael D.
;    APPLICANT:  Shelton, David L.
;    APPLICANT:  Wong, Wai Lee Tan
;    TITLE OF INVENTION:  KINASE RECEPTOR ACTIVATION ASSAY
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,816A
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286305
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/170558
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/157563
;      FILING DATE:  23-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  854C1P1C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2820 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-440-816A-4

  Query Match             90.6%;  Score 15.4;  DB 2;  Length 2820;
  Best Local Similarity   94.1%;  Pred. No. 40;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db    2643 GGGAGCCCCAGCAACGC 2659


RESULT 4
US-09-417-381A-4
; Sequence 4, Application US/09417381A
; Patent No. 6287784
; GENERAL INFORMATION:
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Mark, Melanie R.
;  APPLICANT:  Sadick, Michael D.
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Wong, Wai Lee Tan
;  TITLE OF INVENTION:  KINASE RECEPTOR ACTIVATION ASSAY
;  FILE REFERENCE:  P0854C1P2C1
;  CURRENT APPLICATION NUMBER: US/09/417,381A
```

```
;  CURRENT FILING DATE:  1999-10-13
;  NUMBER OF SEQ ID NOS:  11
; SEQ ID NO 4
;   LENGTH: 2820
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-417-381A-4

  Query Match             90.6%;  Score 15.4;  DB 4;  Length 2820;
  Best Local Similarity   94.1%;  Pred. No. 40;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db    2643 GGGAGCCCCAGCAACGC 2659


RESULT 5
US-08-943-731-2/c
; Sequence 2, Application US/08943731
; Patent No. 6265157
;  GENERAL INFORMATION:
;    APPLICANT:  PROCKOP, DARWIN J.
;    APPLICANT:  SPOTILA, LORETTA D.
;    APPLICANT:  DELTAS, CONSTANTINOS D.
;    APPLICANT:  SEREDA, LARISA
;    APPLICANT:  LARSON, ANDREA W.
;    APPLICANT:  PACK, MICHAEL
;    APPLICANT:  COLIGE, ALAIN
;    APPLICANT:  EARLY, JAMES
;    APPLICANT:  KORKKO, JARMO
;    APPLICANT:  ALA-KOKKO, LEENA, et al.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR DETECTING
;    TITLE OF INVENTION:  ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
;    NUMBER OF SEQUENCES:  666
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PANITCH SCHWARZE JACOBS & NADEL, P.C.
;      STREET:  ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
;      STREET:  FLR.
;      CITY:  PHILADELPHIA
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-7086
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,731
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,322
;      FILING DATE:  14-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/803,628
;      FILING DATE:  03-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOYLE LEARY Ph.D., KATHRYN
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  9598-27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-965-1284
;      TELEFAX:  215-567-2991
;      TELEX:  831-494
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38682 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-943-731-2

  Query Match             90.6%;  Score 15.4;  DB 4;  Length 38682;
  Best Local Similarity   94.1%;  Pred. No. 48;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db   17119 GGGAGCCCCAGCAACGC 17103


RESULT 6
US-09-055-113-2
; Sequence 2, Application US/09055113
; Patent No. 6265565
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Lal, Preeti
;    TITLE OF INVENTION:  PROSTATE-ASSOCIATED ETS PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,113
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0501 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1905 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSTUT12
;      CLONE:  1813005
US-09-055-113-2

  Query Match             88.2%;  Score 15;  DB 4;  Length 1905;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATG 16
           |||||||||||||||
Db     878 GGAGCCCCAGCAATG 892


RESULT 7
US-09-016-000-10/c
; Sequence 10, Application US/09016000
; Patent No. 5962232
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    TITLE OF INVENTION:  PROTEIN KINASE MOLECULES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,000
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0465 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1929 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  MMLR1DT01
;      CLONE:  472480
US-09-016-000-10

  Query Match             88.2%;  Score 15;  DB 2;  Length 1929;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATG 16
           |||||||||||||||
Db      40 GGAGCCCCAGCAATG 26


RESULT 8
US-08-468-551-9
; Sequence 9, Application US/08468551
; Patent No. 5874212
;  GENERAL INFORMATION:
;    APPLICANT:  Prockop, Darwin J.
;    APPLICANT:  Rock, Matthew J.
;    APPLICANT:  Ganguly, Arupa
;    TITLE OF INVENTION:  DETECTION OF SINGLE BASE MUTATIONS AND
;    TITLE OF INVENTION:  OTHER VARIATIONS IN DOUBLE STRANDED DNA BY
;    TITLE OF INVENTION:  CONFORMATION-SENSITIVE CELL ELECTROPHORESIS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PANITCH SCHWARZE JACOBS & NADEL, P.C.
;      STREET:  ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
;      STREET:  FLOOR
;      CITY:  PHILADELPHIA
;      STATE:  PENNSYLVANIA
;      COUNTRY:  UNITED STATES
;      ZIP:  19103-7086
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,551
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Doyle Leary Ph.D., Kathryn
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  9855-5U1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-965-1284
;      TELEFAX:  215-567-2991
;      TELEX:  831-494
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-468-551-9

  Query Match             84.7%;  Score 14.4;  DB 2;  Length 23;
  Best Local Similarity   93.8%;  Pred. No. 92;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           ||||| ||||||||||
Db       8 GGAGCACCAGCAATGC 23


RESULT 9
US-08-717-294-104/c
; Sequence 104, Application US/08717294
; Patent No. 6114148
;  GENERAL INFORMATION:
;    APPLICANT:  SEED, BRIAN
;    APPLICANT:  HAAS, JURGEN
;    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  110
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Clark & Elbing LLP
;      STREET:  176 Federal Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,294
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elbing, Karen L
;      REGISTRATION NUMBER:  35,238
;      REFERENCE/DOCKET NUMBER:  00786/345001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-428-0200
```

```
;       TELEFAX:  617-428-7045
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  104:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  104 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other
US-08-717-294-104

  Query Match             84.7%;  Score 14.4;  DB 3;  Length 104;
  Best Local Similarity   93.8%;  Pred. No. 1e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           |||||||||||||| ||
Db      98 GGAGCCCCAGCAAGGC 83


RESULT 10
US-08-717-294-42
; Sequence 42, Application US/08717294
; Patent No. 6114148
;  GENERAL INFORMATION:
;     APPLICANT:  SEED, BRIAN
;     APPLICANT:  HAAS, JURGEN
;     TITLE OF INVENTION:  HIGH LEVEL EXPRESSION OF
;     TITLE OF INVENTION:  PROTEINS
;     NUMBER OF SEQUENCES:  110
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Clark & Elbing LLP
;       STREET:  176 Federal Street
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02110
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/717,294
;       FILING DATE:  20-SEP-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Elbing, Karen L
;       REGISTRATION NUMBER:  35,238
;       REFERENCE/DOCKET NUMBER:  00786/345001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-428-0200
;       TELEFAX:  617-428-7045
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  42:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4451 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-717-294-42

  Query Match             84.7%;  Score 14.4;  DB 3;  Length 4451;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           |||||||||||||| ||
     4051 GGAGCCCCAGCAAGGC 4066


RESULT 11
US-09-453-702B-162
; Sequence 162, Application US/09453702B
; Patent No. 6365723
;  GENERAL INFORMATION:
;       APPLICANT: Blattner, Frederick R.
;                  Burland,     Valerie
;                  Perna,  Nicole T.
;                  Plunkett, Guy
;                  Welch,  Rod
;       TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Quarles & Brady
;            STREET: 1 South Pinckney Street
;            CITY: Madison
;            STATE: WI
;            COUNTRY: US
;            ZIP: 53701-2113
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Word Perfect 8.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/453,702B
;            FILING DATE: 03-Dec-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/110,955
;            FILING DATE: 04-DEC-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Seay, Nicholas J.
;            REGISTRATION NUMBER: 27386
;            REFERENCE/DOCKET NUMBER: 960296.95017
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (608) 251-5000
;            TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 162:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 9179
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA  (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-09-453-702B-162

  Query Match             84.7%;  Score 14.4;  DB 4;  Length 9179;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           | ||||||||||||||
Db    2750 GCAGCCCCAGCAATGC 2765


RESULT 12
US-08-611-280-4/c
; Sequence 4, Application US/08611280
; Patent No. 5891666
;  GENERAL INFORMATION:
;    APPLICANT:  Matsuyama, Toshifumi
;    APPLICANT:  Grossman, Alex
;    APPLICANT:  Richardson, Christopher D.
;    TITLE OF INVENTION:  NOVEL GENES ENCODING LSIRF POLYPEPTIDES
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Canada Inc.
;      STREET:  6733 Mississauga Road, Suite 303
;      CITY:  Mississauga
```

```
;       STATE:  Ontario
;       COUNTRY:  Canada
;       ZIP:  L5N 6JB
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/611,280
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oleski, Nancy A.
;       REGISTRATION NUMBER:  34,688
;       REFERENCE/DOCKET NUMBER:  A-338A
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12537 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-611-280-4

  Query Match             84.7%;  Score 14.4;  DB 2;  Length 12537;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           ||||||||| ||||||
Db    7199 GGAGCCCCAACAATGC 7184


RESULT 13
US-09-195-940-4/c
; Sequence 4, Application US/09195940
; Patent No. 6258935
;  GENERAL INFORMATION:
;    APPLICANT:  Matsuyama, Toshifumi
;    APPLICANT:  Grossman, Alex
;    APPLICANT:  Richardson, Christopher D.
;    TITLE OF INVENTION:  NOVEL GENES ENCODING LSIRF POLYPEPTIDES
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Canada Inc.
;      STREET:  6733 Mississauga Road, Suite 303
;      CITY:  Mississauga
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  L5N 6JB
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/195,940
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/611,280
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oleski, Nancy A.
;      REGISTRATION NUMBER:  34,688
;      REFERENCE/DOCKET NUMBER:  A-338A
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12537 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-195-940-4

  Query Match             84.7%;  Score 14.4;  DB 4;  Length 12537;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           ||||||||| ||||||
Db    7199 GGAGCCCCAACAATGC 7184


RESULT 14
US-09-562-466-4/c
; Sequence 4, Application US/09562466
; Patent No. 6369202
;  GENERAL INFORMATION:
;       APPLICANT: Matsuyama, Toshifumi
;                  Grossman, Alex
;                  Richardson, Christopher D.
;       TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Amgen Canada Inc.
;            STREET: 6733 Mississauga Road, Suite 303
;            CITY: Mississauga
;            STATE: Ontario
;            COUNTRY: Canada
;            ZIP: L5N 6JB
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/562,466
;            FILING DATE: 01-May-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/195,940
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Oleski, Nancy A.
;            REGISTRATION NUMBER: 34,688
;            REFERENCE/DOCKET NUMBER: A-338A
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 12537 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-562-466-4

  Query Match             84.7%;  Score 14.4;  DB 4;  Length 12537;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGAGCCCCAGCAATGC 17
           ||||||||| ||||||
Db    7199 GGAGCCCCAACAATGC 7184


RESULT 15
US-08-822-774-18
; Sequence 18, Application US/08822774
; Patent No. 6183997
;  GENERAL INFORMATION:
;    APPLICANT:  HOGREFE, Holly
;    TITLE OF INVENTION:  Polymerase Enhancing Factor (PEF)
;    TITLE OF INVENTION:  Extracts, PEF Protein Complexes, Isolated PEF Proteins,
;    TITLE OF INVENTION:  and Methods for Purifying and Identifying Same
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:54:52 ; Search time 142.611 Seconds
(without alignments)
268.450 Million cell updates/sec

Title: US-09-853-688-35
Perfect score: 17
Sequence: 1 gggagccccagcaatgc 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17	100.0	17	24	AAS18882	Growth hormone 1 g
	2	17	100.0	3700	24	AAS18886	Growth hormone 1 g
c	3	15.4	90.6	354	21	AAC30859	Human secreted pro
	4	15.4	90.6	468	22	ABA57641	Human foetal liver
	5	15.4	90.6	468	22	ABA27062	Probe #5528 for ge
	6	15.4	90.6	468	22	AAK05695	Human brain expres
	7	15.4	90.6	468	22	AAK31310	Human bone marrow
	8	15.4	90.6	468	22	AAI15629	Probe #5562 for ge
	9	15.4	90.6	468	22	AAI37207	Probe #5893 used t
	10	15.4	90.6	468	24	ABS06062	Human genome-deriv
	11	15.4	90.6	710	24	ABQ55786	Human ovarian anti
	12	15.4	90.6	711	20	AAZ15243	Human gene express
	13	15.4	90.6	711	20	AAX98799	Human validated ca
	14	15.4	90.6	785	22	AAI94530	Human neuroblastom
	15	15.4	90.6	805	18	AAT87832	Human collagen 1A2
	16	15.4	90.6	852	18	AAT80192	Human collagen 1A2
	17	15.4	90.6	1746	23	AAS84464	DNA encoding novel
	18	15.4	90.6	2383	22	AAK94277	Human full-length
	19	15.4	90.6	2820	16	AAT51456	gD.trkA fusion use
c	20	15.4	90.6	3120	21	AAA12510	cDNA encoding a hu
	21	15.4	90.6	3409	23	ABL29795	Drosophila melanog
	22	15.4	90.6	4556	23	AAS70098	DNA encoding novel
c	23	15.4	90.6	5086	24	ABN97452	Gene #3950 used to
c	24	15.4	90.6	5086	24	ABK64850	Human benign prost
c	25	15.4	90.6	5086	24	ABL92095	Human Tumour Endot
c	26	15.4	90.6	5086	24	ABL92105	Human Tumour Endot
c	27	15.4	90.6	5086	24	ABL92121	Human Tumour Endot
c	28	15.4	90.6	5086	24	ABL62185	Colon adenocarcino
c	29	15.4	90.6	5086	24	ABL62730	Colon adenocarcino
c	30	15.4	90.6	5145	21	AAC98196	Human colon cancer
c	31	15.4	90.6	5416	22	ABA83119	Collagen type I al
c	32	15.4	90.6	5416	24	ABN95596	Gene #2094 used to
c	33	15.4	90.6	5416	24	ABK64492	Human benign prost
c	34	15.4	90.6	5416	24	ABL62092	Colon adenocarcino
c	35	15.4	90.6	5416	24	ABL62655	Colon adenocarcino
c	36	15.4	90.6	5416	24	ABL65476	Lung cancer relate
c	37	15.4	90.6	5564	23	ABV24693	Human prostate exp
	38	15.4	90.6	8778	23	ABL29794	Drosophila melanog
c	39	15.4	90.6	19628	22	AAK71839	Human immune/haema
c	40	15.4	90.6	25576	22	AAK69089	Human immune/haema
c	41	15.4	90.6	38682	22	AAS21770	Human gene for col
	42	15	88.2	392	21	AAH30613	Human colon cancer
	43	15	88.2	1041	22	AAH68387	C glutamicum codin
	44	15	88.2	1087	21	AAF21828	Human breast and o
	45	15	88.2	1089	22	AAS32204	Human DNA repair a

ALIGNMENTS

RESULT 1
AAS18882
ID AAS18882 standard; DNA; 17 BP.
XX
AC AAS18882;
XX
DT 12-MAR-2002 (first entry)
XX
DE Growth hormone 1 gene (GH1) specific fragment, PCR primer GH1F.
XX
KW Growth hormone 1; GH1; osteopathic; gene therapy; protein therapy;
KW diabetes; obesity; infection; acromegaly; gigantism; sodium retention;
KW water retention; metabolic syndrome; mood disorder; sleep disorder;
KW Growth hormone dysfunction; familial growth hormone deficiency;
KW short stature; pituitary storage defect; human; PCR primer; GH1F; ss.
XX
OS Homo sapiens.
XX
PN WO200185993-A2.
XX
PD 15-NOV-2001.
XX
PF 14-MAY-2001; 2001WO-GB02126.
XX
PR 12-MAY-2000; 2000GB-0011459.
PR 14-JUL-2000; 2000EP-0306004.
XX
PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
XX
PI Cooper DN, Procter AM, Gregory J, Millar DS;
XX
DR WPI; 2002-089798/12.

```
XX
PT   Detecting growth hormone variants (GH1), useful in screening patients
PT   for growth hormone irregularities, comprises comparing the nucleotide
PT   sequence of a GH1 gene from a test sample with that of a standard
PT   sequence of the human GH1 -
XX
PS   Example 2; Page 39; 95pp; English.
XX
CC   The invention described a method of detecting variation in growth hormone
CC   1 (GH1), and therefore GH dysfunction in an individual. The method
CC   comprises comparing the nucleotide sequence of GH1 gene obtained from the
CC   test sample with a standard human GH1 gene sequence, in order to identify
CC   variation (GH1 variant). The method is useful in screening patients for
CC   growth hormone irregularities or producing variant proteins for treating
CC   irregularities, and for the early detection and appropriate clinical
CC   management of familial GH deficiency. The GH1 variants are useful in
CC   therapeutic, diagnostic or detection method, particularly for determining
CC   binding defects and susceptibility to a disease such as diabetes, obesity
CC   or infection; for treating acromegaly or gigantism conditions associated
CC   with lactogenic, diabetogenic, lipolytic and protein anabolic effects,
CC   conditions associated with sodium and water retention, metabolic
CC   syndromes, mood and sleep disorders; diagnosing GH dysfunction and
CC   determining pituitary storage defects. The GH1 variants are especially
CC   useful in gene therapy or protein therapy. The GH1 or GH variant may also
CC   be used in the preparation of a medicament, diagnostics composition or
CC   kit, or detection kit. The method has the advantage of: expanding the
CC   know spectrum of GH1 gene mutations; evaluating the role of GH1 gene
CC   mutations in the etiology of short stature; identifying of the mode of
CC   inheritance of novel lesions; evaluation the effects of GH1 mutations on
CC   the structure and function of the GH molecule and development of rapid
CC   diagnostic tests for inherited GH deficiency. This sequence is the GH1
CC   PCR primer, GH1F, used with GH1R (AAS18883) to amplify a GH1-specific
CC   fragment, described in the method of the invention.
XX
SQ   Sequence 17 BP; 4 A; 6 C; 6 G; 1 T; 0 other;

  Query Match             100.0%;  Score 17;  DB 24;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 46;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||||
Db       1 GGGAGCCCCAGCAATGC 17


RESULT 2
AAS18886
ID   AAS18886 standard; DNA; 3700 BP.
XX
AC   AAS18886;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Growth hormone 1 gene (GH1), reference sequence.
XX
KW   Growth hormone 1; GH1; osteopathic; gene therapy; protein therapy;
KW   diabetes; obesity; infection; acromegaly; gigantism; sodium retention;
KW   water retention; metabolic syndrome; mood disorder; sleep disorder;
KW   Growth hormone dysfunction; familial growth hormone deficiency;
KW   short stature; pituitary storage defect; human; chromosome 17q23; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     763..1100
FT                   /*tag= a
FT   CDS             763..2230
FT                   /*tag= b
FT                   /product= "GH1"
FT                   /note= "Growth hormone 1"
FT   variation       replace(124,G)
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(128,T)
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(134,G)
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(135,T)
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(135,C)
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(136,G)
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(141,G)
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(179,C)
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(188,T)
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(218,A)
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(226,G)
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(230,C)
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(234,C)
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(236,C)
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(249,G)
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(281,C)
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(284,A)
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(284,C)
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(286,C)
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(303,C)
FT                   /*tag= v
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(313,A)
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(508)
FT                   /*tag= x
FT                   /note= "Deletion of base A"
FT   variation       replace(519,T)
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(524,A)
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(558,G)
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(565,G)
```

```
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(573,G)
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(580,G)
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(585,G)
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(586,T)
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(620,A)
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(622,G)
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(649,C)
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(665,C)
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(670,G)
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(676,A)
FT                   /*tag= al
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(685,A)
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            701..772
FT                   /*tag= an
FT                   /number= 1
FT   intron          773..1032
FT                   /*tag= ao
FT                   /number= 1
FT   variation       replace(836,C)
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(839,C)
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(879,G)
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(883,A)
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(901,C)
FT                   /*tag= at
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1010,T)
FT                   /*tag= au
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            1033..1193
FT                   /*tag= av
FT                   /number= 2
FT   variation       replace(1097,A)
FT                   /*tag= aw
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1101,T)
FT                   /*tag= ax
FT                   /standard_name= "Single nucleotide polymorphism"
FT   mat_peptide     1101..2227
FT                   /*tag= ay
FT                   /label= "mature_GH1"
FT   variation       replace(1114,T)
FT                   /*tag= az
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1169,A)
FT                   /*tag= ba
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1182,T)
FT                   /*tag= bb
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1189,G)
FT                   /*tag= bc
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1193,G)
FT                   /*tag= bd
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          1194..1402
FT                   /*tag= be
FT                   /number= 2
FT   variation       replace(1196,G)
FT                   /*tag= bf
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1196,C)
FT                   /*tag= bg
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1208,C)
FT                   /*tag= bh
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(1212,T)
FT                   /*tag= bi
FT                   /standard_name= "Single nucleotide polymorphism"

  Query Match             100.0%;  Score 17;  DB 24;  Length 3700;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||||
Db      86 GGGAGCCCCAGCAATGC 102


RESULT 3
AAC30859/c
ID   AAC30859 standard; cDNA; 354 BP.
XX
AC   AAC30859;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 34934.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 34934; 71pp + CD-ROM; English.
XX
```

```
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 354 BP; 90 A; 89 C; 79 G; 96 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 21;  Length 354;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||  ||||||
Db     148 GGGAGCCCCACCAATGC 132


RESULT 4
ABA57641
ID   ABA57641 standard; DNA; 468 BP.
XX
AC   ABA57641;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #5946.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 5946; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
GC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 5
ABA27062
ID   ABA27062 standard; DNA; 468 BP.
XX
AC   ABA27062;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #5528 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 5528; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 6
AAK05695
ID   AAK05695 standard; DNA; 468 BP.
XX
AC   AAK05695;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 5686.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 5686; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 7
AAK31310
ID   AAK31310 standard; DNA; 468 BP.
XX
AC   AAK31310;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 5867.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 5867; 658pp + Sequence Listing; English..
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 8
AAI15629
ID   AAI15629 standard; DNA; 468 BP.
XX
AC   AAI15629;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #5562 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
```

```
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 5562; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 9
AA137207
ID   AA137207 standard; DNA; 468 BP.
XX
AC   AA137207;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #5893 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes usetul tor
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 5893; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 10
ABS06062
ID   ABS06062 standard; DNA; 468 BP.
XX
AC   ABS06062;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 6053.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 1; SEQ ID No 6053; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
```

```
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 468 BP; 113 A; 116 C; 146 G; 93 T; 0 other;

  Query Match             90.6%;  Score 15.4;  DB 24;  Length 468;
  Best Local Similarity   94.1%;  Pred. No. 3.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| |
Db     416 GGGAGCCCCAGCAATTC 432


RESULT 11
ABQ55786
ID   ABQ55786 standard; cDNA; 710 BP.
XX
AC   ABQ55786;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HOPKO61 cDNA, SEQ ID NO:1666.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   P-PSDB; ABP42709.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 1; SEQ ID No 1666; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents cDNA encoding a human ovarian antigen of the
CC   invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 710 BP; 113 A; 244 C; 222 G; 125 T; 6 other;

  Query Match             90.6%;  Score 15.4;  DB 24;  Length 710;
  Best Local Similarity   94.1%;  Pred. No. 3.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           |||||||||| |||||
Db     474 GGGAGCCCCAGAAATGC 490


RESULT 12
AAZ15243
ID   AAZ15243 standard; cDNA; 711 BP.
XX
AC   AAZ15243;
XX
DT   12-OCT-1999  (first entry)
XX
DE   Human gene expression product cDNA sequence SEQ ID NO:2712.
XX
```

```
KW   Human; gene; gene expression product; diagnosis; therapy; probe;
KW   detection; mapping; tissue typing; profiling; forensic; cancer;
KW   genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9938972-A2.
XX
PD   05-AUG-1999.
XX
PF   28-JAN-1999;   99WO-US01619.
XX
PR   03-APR-1998;   98US-0080666.
PR   28-JAN-1998;   98US-0072910.
PR   24-FEB-1998;   98US-0075954.
PR   31-MAR-1998;   98US-0080114.
PR   03-APR-1998;   98US-0080515.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Crkvenjakov R,  Dickson M,  Drmanac R,  Drmanac S;
PI   Escobedo J,  Garcia PD,  Garcia V,  Giese K,  Innis MA;
PI   Jones WL,  Kassam A,  Kennedy GC,  Kita D,  Labat I;
PI   Lamson G,  Leshkowitz D,  Pot D,  Randazzo F,  Reinhard C;
PI   Stache-Crain B,  Sudduth-Klinger J,  Williams LT;
XX
DR   WPI; 1999-494092/41.
XX
PT   Novel human genes and their expression products which are
PT   differentially expressed in different cell types
XX
PS   Claim 1; Page 1317; 2479pp; English.
XX
CC   The present invention describes a library of human polynucleotides
CC   comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC   a method of detecting differentially expressed genes correlated with the
CC   cancerous state of a mammalian cell, comprising detecting at least one
CC   differentially expressed gene product in a test sample from a cell
CC   suspected of being cancerous, where the gene product is encoded by one
CC   of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC   polynucleotides can be used as a source of primers and probes, which can
CC   be used for a variety of purpose, e.g. detection of expression levels,
CC   mapping, tissue typing or profiling, forensics, genetic analysis and
CC   detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC   can be used for raising antibodies for experimental, diagnostic and
CC   therapeutic purposes. The polynucleotides may also be used to construct
CC   arrays for diagnostics (which may be used to determine function of an
CC   encoded protein); and to detect differences in expression levels between
CC   two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC   identify a genetic predisposition or susceptibility to a disease such as
CC   cancer). The polynucleotides of the invention are especially used in the
CC   diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC   and lung cancer. The polynucleotides can also be used to screen for
CC   peptide analogues and antagonists.
XX
SQ   Sequence 711 BP; 137 A; 214 C; 194 G; 159 T; 7 other;

  Query Match             90.6%;  Score 15.4;  DB 20;  Length 711;
  Best Local Similarity   94.1%;  Pred. No. 3.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db     478 GGGAGCCCCAGCAACGC 494


RESULT 13
AAX98799
ID   AAX98799 standard; cDNA; 711 BP.
XX
AC   AAX98799;
XX
DT   24-SEP-1999  (first entry)
XX
DE   Human validated cancer cell derived cDNA #121.
XX
KW   Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
KW   integral membrane protein; aspartyl protease; GATA family; wnt family;
KW   transcription factor; G-protein alpha subunit; protein phosphatase;
KW   phorbolester binding protein; diacylglycerol binding protein; trypsin;
KW   protein kinase; tyrosine phosphatase; developmental signalling protein;
KW   WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;
KW   detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
KW   Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
KW   leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
KW   prostate; ss.
XX
OS   Homo sapiens.
XX
PN   WO9933982-A2.
XX
PD   08-JUL-1999.
XX
PF   22-DEC-1998;   98WO-US27610.
XX
PR   21-DEC-1998;   98US-0217471.
PR   23-DEC-1997;   97US-0068755.
PR   03-APR-1998;   98US-0080664.
PR   21-OCT-1998;   98US-0105234.
PR   27-OCT-1998;   98US-0105877.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Crkvenjakov R,  Dickson M,  Drmanac R,  Drmanac S;
PI   Escobedo J,  Garcia PD,  Garcia V,  Giese K,  Innis MA;
PI   Jones LW,  Kassam A,  Kennedy GC,  Kita D,  Labat I;
PI   Lamson G,  Leshkowitz D,  Pot D,  Randazzo F,  Reinhard C;
PI   Stache-Crain B,  Sudduth-Klinger J,  Williams LT;
XX
DR   WPI; 1999-430243/36.
XX
PT   New isolated human polynucleotides
XX
PS   Claim 1; Page 462; 591pp; English.
XX
CC   This invention describes novel isolated human polynucleotides obtained
CC   by screening for differential expression in colon cancer, breast cancer
CC   and lung cancer cell lines. The polynucleotides of the invention are
CC   represented in AAX98275-X99118 and encode polypeptides of protein
CC   families selected from 4 transmembrane segments integral membrane
CC   proteins, 7 transmembrane receptors, ATPases associated with various
CC   cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of
CC   transcription factors, G-protein alpha subunit, phorbolesters or
CC   diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,
CC   protein tyrosine phosphatase, trypsin, wnt family of developmental
CC   signalling proteins and WW/rsp5/WWP domain containing proteins. The
CC   encoded polypeptides also have a functional domain selected from Ank
CC   repeat, basic region plus leucine zipper transcription factors,
CC   bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger
CC   (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease
CC   domain. The polynucleotides encode polypeptides with similarity to known
CC   protein families and are predicted to have similar properties. The novel
CC   polynucleotides can be used to develop products for use as therapeutic
CC   agents and in forensics, genetic analysis, mapping and diagnostic
CC   applications. In particular, the product can be used for the detection
CC   and management of cancers. They can be used for treating e.g. cervical
CC   cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas,
CC   retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic
CC   myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and
CC   myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric
CC   hereditary ectodermal dysplasia, congenital alveolar dysplasia,
CC   epithelial dysplasia of the cervix, fibrous dysplasia of bone, and
CC   mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,
CC   prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of
```

```
CC   the skin.
XX
SQ   Sequence 711 BP; 137 A; 214 C; 194 G; 159 T; 7 other;

  Query Match             90.6%;  Score 15.4;  DB 20;  Length 711;
  Best Local Similarity   94.1%;  Pred. No. 3.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db     478 GGGAGCCCCAGCAACGC 494


RESULT 14
AA194530
ID   AA194530 standard; cDNA; 785 BP.
XX
AC   AA194530;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Human neuroblastoma expressed polynucleotide SEQ ID NO 605.
XX
KW   Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200166719-A1.
XX
PD   13-SEP-2001.
XX
PF   02-MAR-2001; 2001WO-JP01629.
XX
PR   07-MAR-2000; 2000JP-0159195.
XX
PA   (CHIB-) CHIBA PREFECTURE.
PA   (HISM ) HISAMITSU PHARM CO LTD.
XX
PI   Nakagawara A;
XX
DR   WPI; 2001-565584/63.
XX
PT   Nucleic acids originating in gene expressed in human neuroblastoma,
PT   useful as probe or primer in diagnosing prognosis of human
PT   neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT   for anti-cancer agents -
XX
PS   Claim 1; Page 484-485; 2979pp; Japanese.
XX
CC   The invention relates to novel genes (AAI93926-AAI97963) expressed in
CC   human neuroblastoma. The nucleic acids are applicable as a probe or
CC   primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC   susceptibility indicators or tumour markers for anti-cancer agents. The
CC   gene information for diagnosing prognosis is related to factors similar
CC   to that for N-myc and TrkA genes.
XX
SQ   Sequence 785 BP; 179 A; 261 C; 189 G; 135 T; 21 other;

  Query Match             90.6%;  Score 15.4;  DB 22;  Length 785;
  Best Local Similarity   94.1%;  Pred. No. 3.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           || |||||||||||||||
Db     359 GGAAGCCCCAGCAATGC 375


RESULT 15
AAT87832
ID   AAT87832 standard; cDNA; 805 BP.
XX
AC   AAT87832;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human collagen 1A2 cDNA sequence containing the A polymorphism 907.
XX
KW   Human; collagen 1A2; A polymorphism 907;
KW   endogenous gene suppression; ss.
XX
OS   Homo sapiens.
XX
PN   WO9737014-A1.
XX
PD   09-OCT-1997.
XX
PF   02-APR-1997;   97WO-GB00929.
XX
PR   02-APR-1996;   96GB-0006961.
XX
PA   (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI   Farrar GJ,  Humphries P,  Kenna PF;
XX
DR   WPI; 1997-503100/46.
XX
PT   Endogenous gene suppression and replacement - useful in genetic
PT   disease therapy
XX
PS   Example 4; Page 82; 90pp; English.
XX
CC   The present sequence was used in the development of a novel
CC   strategy for suppressing at least part of the coding region of an
CC   endogenous gene (EG), and replacing the suppressed gene sequence
CC   with a nucleic acid sequence which differs from the EG. The
CC   strategy can be useful when the gene, which is naturally present in
CC   the genome of a patient, contributes to a disease state. Generally,
CC   one allele of the gene will be mutated, i.e. it will possess
CC   alterations in its nucleotide sequence that affect the function or
CC   level of the gene product, e.g. the alteration may result in an
CC   altered protein product from the wild type gene or control of
CC   transcription and processing. Inheritance or somatic acquisition of
CC   such a mutation can give rise, or predispose an individual to a
CC   disease phenotype. However the gene can also be of wild type
CC   phenotype, but contribute to a disease state in another way, so
CC   that suppression would alleviate or improve the disease state or
CC   improve the effectiveness of an administered therapeutic compound.
XX
SQ   Sequence 805 BP; 179 A; 233 C; 162 G; 163 T; 68 other;

  Query Match             90.6%;  Score 15.4;  DB 18;  Length 805;
  Best Local Similarity   94.1%;  Pred. No. 3.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGAGCCCCAGCAATGC 17
           ||||||||||||||| ||
Db      49 GGGAGCCCCAGCAACGC 65


Search completed: January 21, 2003, 16:27:58
Job time : 144.611 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:52:32 ; Search time 1340.03 Seconds
(without alignments)
229.633 Million cell updates/sec

Title: US-09-853-688-36
Perfect score: 19
Sequence: 1 tgtaggaagtctggggtgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19	100.0	629	17	AQ047255	AQ047255 RPCI11-32
	2	17	89.5	478	17	AZ165947	AZ165947 SP_0083_A
c	3	16.4	86.3	223	10	BB148197	BB148197 BB148197
c	4	16.4	86.3	245	10	BB365770	BB365770 BB365770
c	5	16.4	86.3	252	17	CNS04BJ2	AL283223 Tetraodon
c	6	16.4	86.3	302	10	BB041120	BB041120 BB041120
c	7	16.4	86.3	365	17	AQ062287	AQ062287 CIT-HSP-2
	8	16.4	86.3	386	17	AZ739644	AZ739644 RPCI-24-7
c	9	16.4	86.3	570	13	BI064819	BI064819 pgfln.pk0
c	10	16.4	86.3	660	17	AG074600	AG074600 Pan trogl
c	11	16.4	86.3	724	12	BE873650	BE873650 601451618
c	12	16.4	86.3	974	17	CNS0444T	AL273638 Tetraodon
	13	16	84.2	299	14	BQ256354	BQ256354 NISC_jq12
	14	16	84.2	530	17	BH326060	BH326060 CH230-93G
c	15	16	84.2	908	17	AZ206811	AZ206811 SP_0109_A
c	16	16	84.2	1146	14	BM805318	BM805318 AGENCOURT
c	17	15.8	83.2	187	9	AV016693	AV016693 AV016693
c	18	15.8	83.2	216	9	AV249326	AV249326 AV249326
c	19	15.8	83.2	225	9	AV141285	AV141285 AV141285
c	20	15.8	83.2	232	9	AI447750	AI447750 mq81e05.x
	21	15.8	83.2	261	17	AZ704742	AZ704742 RPCI-23-2
	22	15.8	83.2	279	9	AI585575	AI585575 mq81e05.y
c	23	15.8	83.2	279	9	AV211876	AV211876 AV211876
c	24	15.8	83.2	289	10	BB469541	BB469541 BB469541
c	25	15.8	83.2	298	10	BE142089	BE142089 CM2-HT013
	26	15.8	83.2	355	12	BF138757	BF138757 601780353
	27	15.8	83.2	397	9	AA796032	AA796032 vo65a06.r
	28	15.8	83.2	401	9	AA138247	AA138247 mq81e05.r
c	29	15.8	83.2	442	9	AA153692	AA153692 mq59a08.r
c	30	15.8	83.2	444	9	AI645531	AI645531 vt95g11.x
c	31	15.8	83.2	448	17	BH752980	BH752980 SALK_0196
	32	15.8	83.2	484	12	BE918602	BE918602 OV1_8_H10
c	33	15.8	83.2	496	9	AI647277	AI647277 mq59a08.y
c	34	15.8	83.2	508	13	BI811221	BI811221 M001F03 O
c	35	15.8	83.2	521	9	AI721115	AI721115 as73c08.x
	36	15.8	83.2	532	13	BM220909	BM220909 C0947B12-
	37	15.8	83.2	536	14	BQ394880	BQ394880 NISC_ng11
	38	15.8	83.2	542	17	AZ860432	AZ860432 2M0166J24
c	39	15.8	83.2	543	17	AZ883006	AZ883006 RPCI-23-2
c	40	15.8	83.2	554	9	AU148749	AU148749 AU148749
c	41	15.8	83.2	586	10	AV958045	AV958045 AV958045
c	42	15.8	83.2	593	14	BQ745854	BQ745854 UI-M-EX0-
c	43	15.8	83.2	594	14	BP016690	BP016690 BP016690
c	44	15.8	83.2	599	17	AZ358149	AZ358149 1M0100K19
c	45	15.8	83.2	612	17	AZ886074	AZ886074 RPCI-23-1

ALIGNMENTS

RESULT 1
AQ047255/c
LOCUS AQ047255 629 bp DNA linear GSS 14-APR-1999
DEFINITION RPCI11-32I18.TK RPCI-11 Homo sapiens genomic clone RPCI-11-32I18, DNA sequence.
ACCESSION AQ047255
VERSION AQ047255.1 GI:3316182
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 629)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.

FEATURES Location/Qualifiers
source 1. .629
/organism="Homo sapiens"
/db_xref="GDB:7512113"
/db_xref="taxon:9606"
/clone="RPCI-11-32I18"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT 177 a 149 c 129 g 174 t
ORIGIN

Query Match 100.0%; Score 19; DB 17; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGGAAGTCTGGGGTGC 19
||||||||||||||||||||
Db 477 TGTAGGAAGTCTGGGGTGC 459

RESULT 2
AZ165947
LOCUS AZ165947 478 bp DNA linear GSS 29-AUG-2000
DEFINITION SP_0083_A2_G02_T7A Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=83 Col=4 Row=M, DNA sequence.
ACCESSION AZ165947
VERSION AZ165947.1 GI:8335714
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus.
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.
REFERENCE 1 (bases 1 to 478)
AUTHORS Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J, Davidson,E.H. and Hood,L.
TITLE A sea urchin genome project: Sequence scan, virtual map, and additional resources
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
MEDLINE 20402566
COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 83 row: M column: 4
Seq primer: T7
Class: BAC ends
High quality sequence stop: 478.

FEATURES Location/Qualifiers
source 1. .478
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="Plate=83 Col=4 Row=M"
/clone_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli DH10B"

BASE COUNT 150 a 93 c 88 g 140 t 7 others
ORIGIN

Query Match 89.5%; Score 17; DB 17; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TAGGAAGTCTGGGGTGC 19
||||||||||||||||||
Db 53 TAGGAAGTCTGGGGTGC 69

RESULT 3
BB148197/c
LOCUS BB148197 223 bp mRNA linear EST 28-JUN-2000
DEFINITION BB148197 RIKEN full-length enriched, adult female vagina Mus musculus cDNA clone 9930106I05 3', mRNA sequence.
ACCESSION BB148197
VERSION BB148197.1 GI:8803134
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 223)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .223
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="9930106I05"
/clone_lib="RIKEN full-length enriched, adult female vagina"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 50 a 72 c 35 g 66 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 223;
Best Local Similarity 94.4%; Pred. No. 1.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2 GTAGGAAGTCTGGGGTGC 19
           |||||||||||| |||||
Db      77 GTAGGAAGTCTGAGGTGC 60
```

RESULT 4
BB365770/c
LOCUS BB365770 245 bp mRNA linear EST 12-JUL-2000
DEFINITION BB365770 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone C130030N17 3', mRNA sequence.
ACCESSION BB365770
VERSION BB365770.1 GI:9077598
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 245)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1..245
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C130030N17"
/clone_lib="RIKEN full-length enriched, 16 days embryo head"
/sex="mixed"
/tissue_type="head"
/dev_stage="16 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I"
BASE COUNT 76 a 80 c 43 g 46 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 245;
Best Local Similarity 94.4%; Pred. No. 1.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2 GTAGGAAGTCTGGGGTGC 19
           ||||||| ||||||||||
Db      50 GTAGGAAATCTGGGGTGC 33
```

RESULT 5
CNS04BJ2/c
LOCUS CNS04BJ2 252 bp DNA linear GSS 21-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone 097H01 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL283223
VERSION AL283223.1 GI:8021581
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 252)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 252)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished

```
REFERENCE   3  (bases 1 to 252)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .252
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="097H01"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG097CD01SP1-end :
                     PUC-Ori"
BASE COUNT       67 a     58 c     57 g     63 t      7 others
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 17;  Length 252;
  Best Local Similarity   94.4%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           ||| ||||||||||||||
Db     205 TGTGGGAAGTCTGGGGTG 188


RESULT 6
BB041120/c
LOCUS       BB041120                 302 bp    mRNA    linear   EST 23-JUN-2000
DEFINITION  BB041120 RIKEN full-length enriched, 13 days embryo male testis Mus
            musculus cDNA clone 6030453P10 3', mRNA sequence.
ACCESSION   BB041120
VERSION     BB041120.1  GI:8447506
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 302)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .302
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="6030453P10"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     male testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="13 days embryo"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 5.0 and subtraction to Rot = 100.0. Second strand
                     cDNA was prepared with the primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
BASE COUNT       86 a     77 c     51 g     88 t
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 10;  Length 302;
  Best Local Similarity   94.4%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           ||| ||||||||||||||
Db      62 TGTTGGAAGTCTGGGGTG 45


RESULT 7
AQ062287/c
LOCUS       AQ062287                 365 bp    DNA     linear   GSS 31-JUL-1998
DEFINITION  CIT-HSP-2348E20.TF CIT-HSP Homo sapiens genomic clone 2348E20, DNA
            sequence.
ACCESSION   AQ062287
VERSION     AQ062287.1  GI:3364199
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2348E20.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
```

```
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .365
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2348E20"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       97 a      86 c      88 g      94 t
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 17;  Length 365;
  Best Local Similarity   94.4%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           ||||| ||||||||||||
Db     298 TGTAGCAAGTCTGGGGTG 281


RESULT 8
AZ739644
LOCUS       AZ739644                 386 bp    DNA     linear   GSS 25-JAN-2001
DEFINITION  RPCI-24-71H22.TV RPCI-24 Mus musculus genomic clone RPCI-24-71H22,
            DNA sequence.
ACCESSION   AZ739644
VERSION     AZ739644.1  GI:12509933
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 386)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-71H22.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 71  row: H  column: 22
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .386
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-71H22"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
BASE COUNT       96 a      99 c      86 g     105 t
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 17;  Length 386;
  Best Local Similarity   94.4%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           ||||||||| ||||||||
Db     365 TGTAGGAAGACTGGGGTG 382


RESULT 9
BI064819/c
LOCUS       BI064819                 570 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  pgfln.pk002.j5 normalized chicken fat cDNA library Gallus gallus
            cDNA clone pgfln.pk002.j5 5' similar to gi|11525796
            ref|XP_007630.1| KIAA0990 protein [Homo sapiens]G, mRNA sequence.
ACCESSION   BI064819
VERSION     BI064819.1  GI:14472341
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Cogburn,L.A., Morgan,R.W. and Burnside,J.
  TITLE     Chicken ESTs from fat
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="pgfln.pk002.j5"
                     /clone_lib="normalized chicken fat cDNA library"
                     /sex="Male and Female"
                     /tissue_type="fat"
                     /lab_host="E.coli EMDH10B"
                     /note="Vector: pSPORT1"
BASE COUNT      180 a     102 c     115 g     167 t       6 others
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 13;  Length 570;
  Best Local Similarity   94.4%;  Pred. No. 2e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           |||||||||||||| |||
Db     355 TGTAGGAAGTCTGGTGTG 338


RESULT 10
AG074600/c
LOCUS       AG074600                 660 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-066L24.R, genomic survey sequence.
ACCESSION   AG074600
VERSION     AG074600.1  GI:16626402
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-066L24.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
```

```
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 660)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="PTB-066L24.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      159 a    218 c    137 g    146 t
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 17;  Length 660;
  Best Local Similarity   94.4%;  Pred. No. 2.1e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           || |||||||||||||||
Db     321 TGGAGGAAGTCTGGGGTG 304


RESULT 11
BE873650/c
LOCUS       BE873650                724 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601451618F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3855317 5',
            mRNA sequence.
ACCESSION   BE873650
VERSION     BE873650.1  GI:10322426
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9582  row: d  column: 06
            High quality sequence start: 4
            High quality sequence stop: 282.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3855317"
                     /clone_lib="NIH_MGC_65"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies.  "
BASE COUNT      202 a    223 c    137 g    162 t
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 12;  Length 724;
  Best Local Similarity   94.4%;  Pred. No. 2.1e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           |||||||||||||| |||
Db     678 TGTAGGAAGTCTGGCGTG 661


RESULT 12
CNS0444T/c
LOCUS       CNS0444T                974 bp    DNA     linear   GSS 18-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            080L10 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL273638
VERSION     AL273638.1  GI:7995899
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 974)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 974)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 974)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .974
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="080L10"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG080DF05LP1-end : T7"
BASE COUNT      288 a    195 c    198 g    290 t      3 others
ORIGIN

  Query Match             86.3%;  Score 16.4;  DB 17;  Length 974;
  Best Local Similarity   94.4%;  Pred. No. 2.2e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 TGTAGGAAGTCTGGGGTG 18
           ||||||||||||| ||||
Db     357 TGTAGGAAGTCTGAGGTG 340


RESULT 13
BQ256354
LOCUS       BQ256354                 299 bp    mRNA    linear   EST 06-MAY-2002
DEFINITION  NISC_jq12g04.y1 NCI_CGAP_BC2 Mus musculus cDNA clone IMAGE:4113607
            5', mRNA sequence.
ACCESSION   BQ256354
VERSION     BQ256354.1  GI:20457107
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
             MGI:1538759
            Plate: LLAM9369  row: N  column: 8
            Seq primer: M13RPl reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4113607"
                     /clone_lib="NCI_CGAP_BC2"
                     /tissue_type="flow-sorted, common granulocyte-macrophage
                     progenitors"
                     /lab_host="DH10B"
                     /note="Organ: blood; Vector: pAMP1; mRNA made from
                     granulocyte-macrophage progenitors, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       77 a     71 c     96 g     55 t
ORIGIN

  Query Match             84.2%;  Score 16;  DB 14;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||||||||||||||||
Db     120 AGGAAGTCTGGGGTGC 135


RESULT 14
BH326060
LOCUS       BH326060                 530 bp    DNA     linear   GSS 03-DEC-2001
DEFINITION  CH230-93G3.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-93G3, DNA sequence.
ACCESSION   BH326060
VERSION     BH326060.1  GI:17256774
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-93G3.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 93  row: G  column: 3
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Rattus norvegicus"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-93G3"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      123 a    130 c    117 g    160 t
ORIGIN

  Query Match             84.2%;  Score 16;  DB 17;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||||||||||||||||
Db      74 AGGAAGTCTGGGGTGC 89


RESULT 15
AZ206811/c
LOCUS       AZ206811                 908 bp    DNA     linear   GSS 31-AUG-2000
DEFINITION  SP_0109_A1_H02_T7A Strongylocentrotus purpuratus, purple sea urchin
            , sperm genomic BAC library Strongylocentrotus purpuratus genomic
            clone Plate=109 Col=3 Row=O, DNA sequence.
ACCESSION   AZ206811
VERSION     AZ206811.1  GI:8417759
KEYWORDS    GSS.
SOURCE      Strongylocentrotus purpuratus.
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
            Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
            ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J, Davidson,E.H. and
            Hood,L.
  TITLE     A sea urchin genome project: Sequence scan, virtual map, and
            additional resources
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
  MEDLINE   20402566
COMMENT     Contact: Cameron, RA, Davidson, EH, Hood, L
            Division of Biology 156-29
```

California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 109 row: O column: 3
Seq primer: T7
Class: BAC ends
High quality sequence stop: 908.

```
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Strongylocentrotus purpuratus"
                     /db_xref="taxon:7668"
                     /clone="Plate=109 Col=3 Row=O"
                     /clone_lib="Strongylocentrotus purpuratus, purple sea
                     urchin, sperm genomic BAC library"
                     /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
                     DH10B"
BASE COUNT      226 a    194 c    169 g    319 t
ORIGIN

  Query Match             84.2%;  Score 16;  DB 17;  Length 908;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||||||||||||||||
Db     334 AGGAAGTCTGGGGTGC 319


Search completed: January 21, 2003, 15:38:08
Job time : 1345.03 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 14:55:32 ; Search time 34.3056 Seconds
(without alignments)
246.942 Million cell updates/sec

Title: US-09-853-688-36
Perfect score: 19
Sequence: 1 tgtaggaagtctggggtgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	100.0	19	10	US-09-853-688-36	Sequence 36, Appl
c	2	19	100.0	3700	10	US-09-853-688-5	Sequence 5, Appli
c	3	15.8	83.2	232	10	US-09-728-446-1120	Sequence 1120, Ap
c	4	15.4	81.1	170	10	US-09-867-701-388	Sequence 388, App
c	5	15.4	81.1	456	10	US-09-864-761-11292	Sequence 11292, A
	6	15	78.9	392	10	US-09-917-800A-410	Sequence 410, App
c	7	15	78.9	2207	10	US-09-917-800A-1449	Sequence 1449, Ap
c	8	15	78.9	24757	9	US-09-764-868-1347	Sequence 1347, Ap
c	9	14.8	77.9	317	10	US-09-983-965-2137	Sequence 2137, Ap
c	10	14.8	77.9	393	10	US-09-925-300-178	Sequence 178, App
c	11	14.8	77.9	558	10	US-09-864-761-16125	Sequence 16125, A
	12	14.8	77.9	600	10	US-09-878-574-4575	Sequence 4575, Ap
	13	14.8	77.9	1499	10	US-09-800-729-61	Sequence 61, Appl
	14	14.8	77.9	3253	10	US-09-764-855-202	Sequence 202, App
	15	14.8	77.9	3606	10	US-09-764-855-203	Sequence 203, App
c	16	14.8	77.9	7621	10	US-09-954-456-307	Sequence 307, App
c	17	14.8	77.9	7621	10	US-09-954-456-2209	Sequence 2209, Ap
	18	14.8	77.9	10093	10	US-09-764-869-1390	Sequence 1390, Ap
	19	14.4	75.8	459	10	US-09-967-768A-29	Sequence 29, Appl
c	20	14.4	75.8	485	10	US-09-864-761-10605	Sequence 10605, A
c	21	14.4	75.8	1038	10	US-09-804-682-21	Sequence 21, Appl
	22	14.4	75.8	2311	10	US-09-349-015-9	Sequence 9, Appli
c	23	14.4	75.8	3407	10	US-09-971-845-1	Sequence 1, Appli
c	24	14.4	75.8	5935	10	US-09-754-949-8	Sequence 8, Appli
c	25	14.4	75.8	8907	9	US-09-738-626-934	Sequence 934, App
c	26	14.4	75.8	13747	10	US-09-964-824A-287	Sequence 287, App
	27	14.4	75.8	659158	9	US-09-771-208-20	Sequence 20, Appl
	28	14.2	74.7	202	10	US-09-783-590-3921	Sequence 3921, Ap
c	29	14.2	74.7	254	10	US-09-878-574-15524	Sequence 15524, A
c	30	14.2	74.7	344	10	US-09-920-300A-1145	Sequence 1145, Ap
c	31	14.2	74.7	344	12	US-10-033-528-1145	Sequence 1145, Ap
	32	14.2	74.7	491	10	US-09-864-761-5763	Sequence 5763, Ap
	33	14.2	74.7	554	10	US-09-917-800A-346	Sequence 346, App
	34	14.2	74.7	586	10	US-09-739-262-3	Sequence 3, Appli
c	35	14.2	74.7	1874	10	US-09-822-849A-281	Sequence 281, App
c	36	14.2	74.7	1883	10	US-09-822-849A-518	Sequence 518, App
	37	14.2	74.7	2291	10	US-09-883-825-5	Sequence 5, Appli
c	38	14.2	74.7	2476	9	US-09-774-639-31	Sequence 31, Appl
c	39	14.2	74.7	5294	9	US-10-098-841-9	Sequence 9, Appli
	40	14.2	74.7	6393	10	US-09-764-864-1718	Sequence 1718, Ap
c	41	14.2	74.7	6815	10	US-09-764-864-1719	Sequence 1719, Ap
c	42	14.2	74.7	66804	10	US-09-740-041-3	Sequence 3, Appli
	43	14	73.7	282	10	US-09-878-574-7318	Sequence 7318, Ap
c	44	14	73.7	526	10	US-09-864-761-7087	Sequence 7087, Ap
c	45	14	73.7	3348	10	US-09-917-800A-1477	Sequence 1477, Ap

ALIGNMENTS

RESULT 1
US-09-853-688-36
; Sequence 36, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; TITLE OF INVENTION: HUMANS, THE VARIATIONS AND THEIR USES
; FILE REFERENCE: WCM78
; CURRENT APPLICATION NUMBER: US/09/853,688
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-688-36

Query Match 100.0%; Score 19; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGGAAGTCTGGGGTGC 19
||||||||||||||||||||
Db 1 TGTAGGAAGTCTGGGGTGC 19

RESULT 2
US-09-853-688-5/c
; Sequence 5, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; TITLE OF INVENTION: HUMANS, THE VARIATIONS AND THEIR USES

```
;  FILE REFERENCE: WCM78
;  CURRENT APPLICATION NUMBER: US/09/853,688
;  CURRENT FILING DATE:  2001-05-14
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 3700
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-853-688-5

  Query Match             100.0%;  Score 19;  DB 10;  Length 3700;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           |||||||||||||||||||
Db    3316 TGTAGGAAGTCTGGGGTGC 3298


RESULT 3
US-09-728-446-1120/c
; Sequence 1120, Application US/09728446
; Patent No. US20020081668A1
; GENERAL INFORMATION:
;  APPLICANT: Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. US20020081668Alel Murine Polynucleotide Sequences
;  TITLE OF INVENTION:  and Mutant Cells and Mutant Animals Defined Thereby
;  FILE REFERENCE: LEX-0101-USA
;  CURRENT APPLICATION NUMBER: US/09/728,446
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/168,270
;  PRIOR FILING DATE: 1999-12-01
;  NUMBER OF SEQ ID NOS: 1461
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1120
;   LENGTH: 232
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(232)
;   OTHER INFORMATION: n = A,T,C or G
US-09-728-446-1120

  Query Match             83.2%;  Score 15.8;  DB 10;  Length 232;
  Best Local Similarity   89.5%;  Pred. No. 44;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           || |||||||||| |||||
Db     144 TGAAGGAAGTCTGTGGTGC 126


RESULT 4
US-09-867-701-388/c
; Sequence 388, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 388
;   LENGTH: 170
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(170)
;   OTHER INFORMATION: n = A,T,C or G
US-09-867-701-388

  Query Match             81.1%;  Score 15.4;  DB 10;  Length 170;
  Best Local Similarity   94.1%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GTAGGAAGTCTGGGGTG 18
           || ||||||||||||||
Db     115 GTGGGAAGTCTGGGGTG 99


RESULT 5
US-09-864-761-11292/c
; Sequence 11292, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11292
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005342.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.76
US-09-864-761-11292

  Query Match             81.1%;  Score 15.4;  DB 10;  Length 456;
  Best Local Similarity   94.1%;  Pred. No. 73;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGT 17
           ||||||||||| ||||||
Db     125 TGTAGGAAGTTTGGGGT 109


RESULT 6
US-09-917-800A-410
; Sequence 410, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 410
;   LENGTH: 392
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA957202
US-09-917-800A-410

  Query Match             78.9%;  Score 15;  DB 10;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 TAGGAAGTCTGGGGT 17
           |||||||||||||||
Db     190 TAGGAAGTCTGGGGT 204


RESULT 7
US-09-917-800A-1449/c
;
; Sequence 1449, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1449
;   LENGTH: 2207
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 U10357
US-09-917-800A-1449

  Query Match             78.9%;  Score 15;  DB 10;  Length 2207;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 TAGGAAGTCTGGGGT 17
           |||||||||||||||
Db    2036 TAGGAAGTCTGGGGT 2022


RESULT 8
US-09-764-868-1347/c
; Sequence 1347, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PT232
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
;   LENGTH: 24757
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-868-1347

  Query Match             78.9%;  Score 15;  DB 9;  Length 24757;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       5 GGAAGTCTGGGGTGC 19
           ||||||||||||||||
Db   10501 GGAAGTCTGGGGTGC 10487


RESULT 9
US-09-983-965-2137/c
; Sequence 2137, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2137
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 37-LIB3057-014-Q1-K1-B2
US-09-983-965-2137

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 317;
  Best Local Similarity   88.9%;  Pred. No. 1.4e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           || |||||||| |||||||
Db     200 TGAAGGAAGTGTGGGGTG 183


RESULT 10
US-09-925-300-178/c
; Sequence 178, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (214)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-178

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 393;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||| |||| |||||||
Db     224 TGTAGCAAGTNAGGGGTGC 206


RESULT 11
US-09-864-761-16125/c
; Sequence 16125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16125
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010486.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.47
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.46
US-09-864-761-16125

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 558;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GTAGGAAGTCTGGGGTGC 19
           | |||||||| |||||||
Db     362 GCAGGAAGTCAGGGGTGC 345
```

```
RESULT 12
US-09-878-574-4575
; Sequence 4575, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4575
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(600)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3028-033-Q1-B1-H1
US-09-878-574-4575

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 600;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GTAGGAAGTCTGGGGTGC 19
           | ||||||| |||||||||
Db     464 GGAGGAAGCCTGGGGTGC 481


RESULT 13
US-09-800-729-61
; Sequence 61, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
;  APPLICANT:  Ni et al.
;  TITLE OF INVENTION:  32 Human secreted proteins
;  FILE REFERENCE:  PZ044P1
;  CURRENT APPLICATION NUMBER: US/09/800,729
;  CURRENT FILING DATE:  2001-03-08
;  PRIOR APPLICATION NUMBER:  PCT/US00/26013
;  PRIOR FILING DATE:  2000-09-22
;  PRIOR APPLICATION NUMBER:  60/155,709
;  PRIOR FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  217
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 61
;   LENGTH: 1499
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-800-729-61

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 1499;
  Best Local Similarity   88.9%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GTAGGAAGTCTGGGGTGC 19
           |||||||| ||||| |||
Db    1012 GTAGGAAGACTGGGCTGC 1029


RESULT 14
US-09-764-855-202
; Sequence 202, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA110
;  CURRENT APPLICATION NUMBER: US/09/764,855
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 334
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 202
;   LENGTH: 3253
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-855-202

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 3253;
  Best Local Similarity   88.9%;  Pred. No. 1.7e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           || ||||||||| |||||
Db    1624 TGGAGGAAGTCTTGGGTG 1641


RESULT 15
US-09-764-855-203
; Sequence 203, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA110
;  CURRENT APPLICATION NUMBER: US/09/764,855
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 334
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 203
;   LENGTH: 3606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-855-203

  Query Match             77.9%;  Score 14.8;  DB 10;  Length 3606;
  Best Local Similarity   88.9%;  Pred. No. 1.7e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           || ||||||||| |||||
Db    1976 TGGAGGAAGTCTTGGGTG 1993


Search completed: January 21, 2003, 16:29:18
Job time : 37.3056 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 15:09:37 ; Search time 32.7222 Seconds
(without alignments)
178.070 Million cell updates/sec

Title: US-09-853-688-36
Perfect score: 19
Sequence: 1 tgtaggaagtctggggtgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.8	83.2	2246	4	US-09-032-742-3	Sequence 3, Appli
c	2	15.8	83.2	2246	4	US-09-032-742-21	Sequence 21, Appl
c	3	15.8	83.2	2246	4	US-09-032-742-22	Sequence 22, Appl
c	4	15.8	83.2	2246	4	US-09-032-742-24	Sequence 24, Appl
c	5	15.8	83.2	2246	4	US-09-032-742-25	Sequence 25, Appl
c	6	14.8	77.9	3388	4	US-09-141-206-1	Sequence 1, Appli
c	7	14.4	75.8	2511	2	US-08-422-699A-8	Sequence 8, Appli
c	8	14.4	75.8	2511	2	US-08-422-706B-8	Sequence 8, Appli
c	9	14.4	75.8	2726	2	US-08-422-699A-12	Sequence 12, Appl
c	10	14.4	75.8	2726	2	US-08-422-706B-12	Sequence 12, Appl
c	11	14.4	75.8	3182	1	US-08-484-044-11	Sequence 11, Appl
c	12	14.4	75.8	3323	2	US-08-422-699A-10	Sequence 10, Appl
c	13	14.4	75.8	3323	2	US-08-422-706B-10	Sequence 10, Appl
c	14	14.4	75.8	11613	1	US-08-484-044-10	Sequence 10, Appl
c	15	14.4	75.8	20303	1	US-08-370-975B-6	Sequence 6, Appli
c	16	14.4	75.8	26764	1	US-08-370-975B-1	Sequence 1, Appli
	17	14.4	75.8	72928	3	US-09-009-913-1	Sequence 1, Appli
	18	14.2	74.7	94	1	US-08-909-725-1	Sequence 1, Appli
	19	14.2	74.7	586	4	US-09-146-969-3	Sequence 3, Appli
c	20	14.2	74.7	704	4	US-08-998-416-977	Sequence 977, App
	21	14.2	74.7	747	2	US-08-401-530A-1	Sequence 1, Appli
	22	14.2	74.7	747	2	US-08-709-662-1	Sequence 1, Appli
c	23	14.2	74.7	1878	4	US-09-732-025-1	Sequence 1, Appli
	24	14.2	74.7	2291	1	US-07-872-644-5	Sequence 5, Appli
	25	14.2	74.7	2291	1	US-08-297-494-5	Sequence 5, Appli
	26	14.2	74.7	2291	1	US-08-297-510-5	Sequence 5, Appli
	27	14.2	74.7	2291	1	US-08-479-532-5	Sequence 5, Appli
	28	14.2	74.7	2291	1	US-08-455-526-5	Sequence 5, Appli
	29	14.2	74.7	2291	1	US-08-455-525-5	Sequence 5, Appli
	30	14.2	74.7	2291	3	US-09-139-491-5	Sequence 5, Appli
	31	14.2	74.7	2291	5	PCT-US92-03222-5	Sequence 5, Appli
c	32	14.2	74.7	2808	3	US-08-870-126-7	Sequence 7, Appli
c	33	14.2	74.7	2808	4	US-09-445-247-7	Sequence 7, Appli
c	34	14	73.7	3808	2	US-08-916-917-3	Sequence 3, Appli
c	35	14	73.7	3808	2	US-08-972-631-3	Sequence 3, Appli
c	36	14	73.7	3808	2	US-08-972-629-3	Sequence 3, Appli
c	37	14	73.7	3808	2	US-08-972-630-3	Sequence 3, Appli
c	38	14	73.7	3808	2	US-08-672-211-3	Sequence 3, Appli
c	39	14	73.7	3808	3	US-09-225-170-3	Sequence 3, Appli
	40	13.8	72.6	588	4	US-09-385-982-128	Sequence 128, App
	41	13.8	72.6	1646	4	US-09-221-017B-654	Sequence 654, App
c	42	13.8	72.6	2968	4	US-09-813-819-1	Sequence 1, Appli
c	43	13.8	72.6	2968	4	US-09-920-048-1	Sequence 1, Appli
c	44	13.8	72.6	3022	4	US-09-193-562D-33	Sequence 33, Appl
	45	13.8	72.6	11832	2	US-08-416-603-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-032-742-3/c
; Sequence 3, Application US/09032742
; Patent No. 6255089
; GENERAL INFORMATION:
; APPLICANT: Teitler, Milt
; APPLICANT: Herrick-Davis, Katharine
; APPLICANT: Egan, Christina C.
; TITLE OF INVENTION: Constitutively Activated Serotonin
; TITLE OF INVENTION: Receptors
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Laurence Weinberger
; STREET: 882 S. Matlack Street, Suite 103
; STREET: P.O. Box 1663
; CITY: West Chester
; STATE: PA
; COUNTRY: USA
; ZIP: 19380-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,742
; FILING DATE: 27-FEB-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Weinberger, Laurence
; REGISTRATION NUMBER: 27,965
; REFERENCE/DOCKET NUMBER: 3086-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 431-1703
; TELEFAX: (610) 431-4181
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-032-742-3

Query Match 83.2%; Score 15.8; DB 4; Length 2246;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTAGGAAGTCTGGGGTGC 19
||||||||||||| | |||

```
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 2
US-09-032-742-21/c
; Sequence 21, Application US/09032742
; Patent No. 6255089
;  GENERAL INFORMATION:
;    APPLICANT:  Teitler, Milt
;    APPLICANT:  Herrick-Davis, Katharine
;    APPLICANT:  Egan, Christina C.
;    TITLE OF INVENTION:  Constitutively Activated Serotonin
;    TITLE OF INVENTION:  Receptors
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Laurence Weinberger
;      STREET:  882 S. Matlack Street, Suite 103
;      STREET:  P.O. Box 1663
;      CITY:  West Chester
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19380-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,742
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weinberger, Laurence
;      REGISTRATION NUMBER:  27,965
;      REFERENCE/DOCKET NUMBER:  3086-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 431-1703
;      TELEFAX:  (610) 431-4181
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2246 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-032-742-21

  Query Match             83.2%;  Score 15.8;  DB 4;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 3
US-09-032-742-22/c
; Sequence 22, Application US/09032742
; Patent No. 6255089
;  GENERAL INFORMATION:
;    APPLICANT:  Teitler, Milt
;    APPLICANT:  Herrick-Davis, Katharine
;    APPLICANT:  Egan, Christina C.
;    TITLE OF INVENTION:  Constitutively Activated Serotonin
;    TITLE OF INVENTION:  Receptors
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Laurence Weinberger
;      STREET:  882 S. Matlack Street, Suite 103
;      STREET:  P.O. Box 1663
;      CITY:  West Chester
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19380-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,742
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weinberger, Laurence
;      REGISTRATION NUMBER:  27,965
;      REFERENCE/DOCKET NUMBER:  3086-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 431-1703
;      TELEFAX:  (610) 431-4181
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2246 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-032-742-22

  Query Match             83.2%;  Score 15.8;  DB 4;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 4
US-09-032-742-24/c
; Sequence 24, Application US/09032742
; Patent No. 6255089
;  GENERAL INFORMATION:
;    APPLICANT:  Teitler, Milt
;    APPLICANT:  Herrick-Davis, Katharine
;    APPLICANT:  Egan, Christina C.
;    TITLE OF INVENTION:  Constitutively Activated Serotonin
;    TITLE OF INVENTION:  Receptors
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Laurence Weinberger
;      STREET:  882 S. Matlack Street, Suite 103
;      STREET:  P.O. Box 1663
;      CITY:  West Chester
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19380-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,742
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weinberger, Laurence
;      REGISTRATION NUMBER:  27,965
;      REFERENCE/DOCKET NUMBER:  3086-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 431-1703
;      TELEFAX:  (610) 431-4181
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  2246 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-032-742-24

  Query Match             83.2%;  Score 15.8;  DB 4;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 5
US-09-032-742-25/c
; Sequence 25, Application US/09032742
; Patent No. 6255089
;  GENERAL INFORMATION:
;     APPLICANT:  Teitler, Milt
;     APPLICANT:  Herrick-Davis, Katharine
;     APPLICANT:  Egan, Christina C.
;     TITLE OF INVENTION:  Constitutively Activated Serotonin
;     TITLE OF INVENTION:  Receptors
;     NUMBER OF SEQUENCES:  25
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Laurence Weinberger
;       STREET:  882 S. Matlack Street, Suite 103
;       STREET:  P.O. Box 1663
;       CITY:  West Chester
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19380-0053
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/032,742
;       FILING DATE:  27-FEB-1998
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Weinberger, Laurence
;       REGISTRATION NUMBER:  27,965
;       REFERENCE/DOCKET NUMBER:  3086-4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (610) 431-1703
;       TELEFAX:  (610) 431-4181
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2246 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-032-742-25

  Query Match             83.2%;  Score 15.8;  DB 4;  Length 2246;
  Best Local Similarity   89.5%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTGC 19
           ||||||||||||| | |||
Db    2096 TGTAGGAAGTCTGCGCTGC 2078


RESULT 6
US-09-141-206-1/c
; Sequence 1, Application US/09141206
; Patent No. 6187559
; GENERAL INFORMATION:
;  APPLICANT: Steed, Paul M.
;  APPLICANT: Lasala, Daniel J.
;  TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
; Patent No. 6187559
;  FILE REFERENCE: 4-30148/P1/CGC1954/R
;  CURRENT APPLICATION NUMBER: US/09/141,206
;  CURRENT FILING DATE: 1998-08-27
;  EARLIER APPLICATION NUMBER: 60/057,802
;  EARLIER FILING DATE: 1997-08-28
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3388
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (162)..(2963)
;   OTHER INFORMATION: Human PLD2
US-09-141-206-1

  Query Match             77.9%;  Score 14.8;  DB 4;  Length 3388;
  Best Local Similarity   88.9%;  Pred. No. 86;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGTAGGAAGTCTGGGGTG 18
           || | ||||||||||||||
Db    1663 TGAAAGAAGTCTGGGGTG 1646


RESULT 7
US-08-422-699A-8/c
; Sequence 8, Application US/08422699A
; Patent No. 5955265
;  GENERAL INFORMATION:
;     APPLICANT:  Brook, J. David
;     APPLICANT:  Housman, David E.
;     APPLICANT:  Shaw, Duncan J.
;     APPLICANT:  Harley, Helen G.
;     APPLICANT:  Johnson, Keith J.
;     TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;     TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;       STREET:  Two Militia Drive
;       CITY:  Lexington
;       STATE:  Massachusetts
;       COUNTRY:  US
;       ZIP:  02713
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/422,699A
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/422,706
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/023,612
;       FILING DATE:  26-FEB-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/839,255
;       FILING DATE:  20-FEB-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/01545
;       FILING DATE:  19-FEB-1993
;     PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  PCT/GB93/00253
;       FILING DATE:  05-FEB-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB9202485.0
;       FILING DATE:  06-FEB-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Granahan, Patricia
;       REGISTRATION NUMBER:  32,227
;       REFERENCE/DOCKET NUMBER:  MIT-5830A2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-861-6240
;       TELEFAX:  617-861-9540
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2511 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1746
US-08-422-699A-8

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 2511;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db     590 AGGTAGTCTGGGGTGC 575


RESULT 8
US-08-422-706B-8/c
; Sequence 8, Application US/08422706B
; Patent No. 5977333
;  GENERAL INFORMATION:
;    APPLICANT:  Brook, J. David
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Shaw, Duncan J.
;    APPLICANT:  Harley, Helen G.
;    APPLICANT:  Johnson, Keith J.
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;    TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02713
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,706B
;      FILING DATE:  14-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/284,543
;      FILING DATE:  08-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/023,612
;      FILING DATE:  26-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/839,255
;      FILING DATE:  20-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01545
;      FILING DATE:  19-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00253
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB9202485.0
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5830A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1746
US-08-422-706B-8

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 2511;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db     590 AGGTAGTCTGGGGTGC 575


RESULT 9
US-08-422-699A-12/c
; Sequence 12, Application US/08422699A
; Patent No. 5955265
;  GENERAL INFORMATION:
;    APPLICANT:  Brook, J. David
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Shaw, Duncan J.
;    APPLICANT:  Harley, Helen G.
;    APPLICANT:  Johnson, Keith J.
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;    TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02713
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,699A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/422,706
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/023,612
;      FILING DATE:  26-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/839,255
;      FILING DATE:  20-FEB-1992
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01545
;      FILING DATE:  19-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00253
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB9202485.0
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5830A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-422-699A-12

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 2726;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db     840 AGGTAGTCTGGGGTGC 825


RESULT 10
US-08-422-706B-12/c
; Sequence 12, Application US/08422706B
; Patent No. 5977333
;  GENERAL INFORMATION:
;    APPLICANT:  Brook, J. David
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Shaw, Duncan J.
;    APPLICANT:  Harley, Helen G.
;    APPLICANT:  Johnson, Keith J.
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;    TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02713
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,706B
;      FILING DATE:  14-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/284,543
;      FILING DATE:  08-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/023,612
;      FILING DATE:  26-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/839,255
;      FILING DATE:  20-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01545
;      FILING DATE:  19-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00253
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB9202485.0
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5830A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-422-706B-12

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 2726;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db     840 AGGTAGTCTGGGGTGC 825


RESULT 11
US-08-484-044-11/c
; Sequence 11, Application US/08484044
; Patent No. 5552282
;  GENERAL INFORMATION:
;    APPLICANT:  Caskey, C. T.
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Friedman, David L.
;    APPLICANT:  Pizzuti, Antonio
;    APPLICANT:  Fenwick, Raymond G.
;    TITLE OF INVENTION:  Diagnosis of Myotonic Muscular Dystrophy
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski, L.L.P.
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,044
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/019,940
;      FILING DATE:  19-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5443
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5325
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3182 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-484-044-11

  Query Match             75.8%;  Score 14.4;  DB 1;  Length 3182;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db    1288 AGGTAGTCTGGGGTGC 1273


RESULT 12
US-08-422-699A-10/c
; Sequence 10, Application US/08422699A
; Patent No. 5955265
;  GENERAL INFORMATION:
;    APPLICANT:  Brook, J. David
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Shaw, Duncan J.
;    APPLICANT:  Harley, Helen G.
;    APPLICANT:  Johnson, Keith J.
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;    TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02713
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,699A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/422,706
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/023,612
;      FILING DATE:  26-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/839,255
;      FILING DATE:  20-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01545
;      FILING DATE:  19-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00253
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB9202485.0
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5830A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3323 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(518..3323, "")
;      OTHER INFORMATION:  /standard_name= "cDNA 41"
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(769..3323, "")
;      OTHER INFORMATION:  /standard_name= "cDNA 28"
US-08-422-699A-10

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 3323;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db    1375 AGGTAGTCTGGGGTGC 1360


RESULT 13
US-08-422-706B-10/c
; Sequence 10, Application US/08422706B
; Patent No. 5977333
;  GENERAL INFORMATION:
;    APPLICANT:  Brook, J. David
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Shaw, Duncan J.
;    APPLICANT:  Harley, Helen G.
;    APPLICANT:  Johnson, Keith J.
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING THE MYOTONIC
;    TITLE OF INVENTION:  DYSTROPHY GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02713
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,706B
;      FILING DATE:  14-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/284,543
;      FILING DATE:  08-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/023,612
;      FILING DATE:  26-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/839,255
;      FILING DATE:  20-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/01545
;      FILING DATE:  19-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00253
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB9202485.0
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5830A2
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3323 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(518..3323, "")
;      OTHER INFORMATION:  /standard_name= "cDNA 41"
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(769..3323, "")
;      OTHER INFORMATION:  /standard_name= "cDNA 28"
US-08-422-706B-10

  Query Match             75.8%;  Score 14.4;  DB 2;  Length 3323;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db    1375 AGGTAGTCTGGGGTGC 1360


RESULT 14
US-08-484-044-10/c
; Sequence 10, Application US/08484044
; Patent No. 5552282
;  GENERAL INFORMATION:
;    APPLICANT:  Caskey, C. T.
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Friedman, David L.
;    APPLICANT:  Pizzuti, Antonio
;    APPLICANT:  Fenwick, Raymond G.
;    TITLE OF INVENTION:  Diagnosis of Myotonic Muscular Dystrophy
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski, L.L.P.
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,044
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/019,940
;      FILING DATE:  19-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5443
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5325
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11613 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-484-044-10

  Query Match             75.8%;  Score 14.4;  DB 1;  Length 11613;
  Best Local Similarity   93.8%;  Pred. No. 1.6e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AGGAAGTCTGGGGTGC 19
           ||| ||||||||||||
Db    4077 AGGTAGTCTGGGGTGC 4062


RESULT 15
US-08-370-975B-6/c
; Sequence 6, Application US/08370975B
; Patent No. 5622851
;  GENERAL INFORMATION:
;    APPLICANT:  Maley, Frank
;    APPLICANT:  Maley, Gladys F.
;    APPLICANT:  Weiner, Karen X.B.
;    TITLE OF INVENTION:  Human Deoxycytidylate Deaminase Gene
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,975B
;      FILING DATE:  10-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  20894/80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716)263-1636
;      TELEFAX:  (716)263-1600
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20303 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  4q35
US-08-370-975B-6

  Query Match             75.8%;  Score 14.4;  DB 1;  Length 20303;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TAGGAAGTCTGGGGTG 18
           | ||||||||||||||
Db    2205 TGGGAAGTCTGGGGTG 2190


Search completed: January 21, 2003, 16:30:42
Job time : 36.7222 secs
```